

```

CC genes or in urea or ureb which renders them urease-negative or
CC urease-depleted (attenuated strains) can be used as immunogens to
CC protect against H.pylori infection.
SQ Sequence 3560 BP; 1059 A; 699 C; 776 G; 1026 T;

Query Match 75.0%; Score 15; DB 6; Length 3560;
Best Local Similarity 94.1%; Pred. No. 6.18e+00;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 702 tccttttagggaaattc 718
Cp 18 tccttttaggggaattc 2

RESULT 5
ID N80975 standard; DNA; 780 BP.
AC N80975;
DT 12-SEP-1990 (first entry)
Probe PCS.7 for screening for cystic fibrosis-associated RFLP in Hhai
digests of human chromosome 7q 22-31 DNA
KW restriction fragment length polymorphism (RFLP); cosmid CNX.4; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_feature 213
FT /*tag= a
FT /note="polymorphism occurs here, see CC"

PN EP-288299-A.
PD 26-OCT-1988.
PF 22-APR-1988; 303645.
PR 23-APR-1987; GB-009652,
PA (STNA) St Marys Hospital.
PI Williamson R;
DR WPI; 88-301321/43.
PT Region of human genomic DNA from chromosome 7722-31 -
PT used for producing a DNA hybridisation probe for cystic fibrosis
PT risk analysis and diagnosis
PS Disclosure; Page 12; 15pp; English.
CC CNX.4, which contains a putative HpaII tiny fragments (HFF) island. It
CC detects a frequent RFLP (Al 0.56, A2 0.44) in Hhai digests. In 70 out of
CC 71 chromosomes studied, the CF mutation is associated with the PCS.7 A2
CC allele (0.47kb). The polymorphism occurs at tag a (see FT) where the
CC base pair GC provides Hhai site, while the alternative of AT eliminates
CC the site. The probe, which is claimed, can be used in conjunction with
CC haplotyping or carrier exclusion and would enable about two-thirds of the
CC population to be excluded from significant risk.
Sequence 780 BP; 125 A; 261 C; 230 G; 164 T;

Query Match 70.0%; Score 14; DB 1; Length 780;
Best Local Similarity 93.8%; Pred. No. 2.28e+01;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 246 ggagcttcctccaaagg 261
Qy 1 ggaacttcctccaaagg 16

RESULT 6
ID T68000 standard; DNA; 1041 BP.
AC T68000;
DT 16-JUL-1997 (first entry)
DE H. pylori inner membrane protein ORF 06ep10306orf3.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT cds 1..1041
FT /*tag= a
FT /note="no stop codon given"

PN W09640893-A1.

```

```

PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR ) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaard BL;
DR WPI; 97-052306/05.
DR P-PSDB; W20747.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 1; Page 838; 1481pp; English.
CC The present sequence encodes a H. pylori inner membrane protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 1041 BP; 312 A; 200 C; 212 G; 317 T;

Query Match 70.0%; Score 14; DB 29; Length 1041;
Best Local Similarity 93.8%; Pred. No. 2.28e+01;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 46 acttcctctaaaggagg 61
Qy 4 acttcctctaaaggagg 19

RESULT 7
ID T67710 standard; DNA; 1065 BP.
AC T67710;
DT 17-JUL-1997 (first entry)
DE H. pylori cell envelope protein ORF 6828218.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT cds 1..1065
FT /*tag= a
FT /note="no stop codon given"

PN W09640893-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR ) ASTRA AB.
PI Berglindh OT, Smith D, Mellgaard BL;
DR WPI; 97-052306/05.
DR P-PSDB; W20563.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 1; Page 315; 1481pp; English.
CC This sequence encodes a H. pylori cell envelope protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant

```

CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 1065 BP; 279 A; 205 C; 188 G; 393 T;

Query Match 70.0%; Score 14; DB 29; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 2.28e+01;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 950 ctttaggaagttc 963
 |||||
 Cp 15 ctttaggaagttc 2

RESULT 8
 ID T67988 standard; DNA; 1074 BP.

AC T67988;
 DT 16-JUL-1997 (first entry)
 NA H. pylori cell envelope protein ORF 06cp11722orf1.
 CC Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 CC identification; binding compound; bacterium; life cycle; activator;
 KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
 KW diagnosis; ds.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT cds 1..1074
 FT /*tag= a
 FT /note= "no stop codon given"

FN W09640893-A1.
 PD 19-DEC-1996.
 PF 06-JUN-1996; U09122.
 PR 07-JUN-1995; US-487032.
 PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL;
 DT WPI; 97-052306/05.
 DR P-PSDB: W20735.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptides) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PS Claim 1; Page 830; 1481pp; English.
 CC The present sequence encodes a H. pylori cell envelope protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SQ Sequence 1074 BP; 282 A; 205 C; 191 G; 396 T;

Query Match 70.0%; Score 14; DB 29; Length 1074;
 Best Local Similarity 100.0%; Pred. No. 2.28e+01;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 959 ctttaggaagttc 972
 |||||
 Cp 15 ctttaggaagttc 2

RESULT 9
 ID Q43719 standard; DNA; 1404 BP.

AC Q43719;
 DT 25-SEP-1993 (first entry)
 DE Sequence which hybridises specifically to bovine male DNA.
 KW Bovine; embryo; sex selection; ss.

OS Bos primigenius, strain Holstein.
 PN EP-546762-A.
 PD 16-JUN-1993.
 PF 02-DEC-1992; 311013.
 PR 13-DEC-1991; JP-352032.
 PA (ITOH-) ITOHAM FOODS INC.
 PI Itagaki Y, Kudo T, Nakamura T, Sato S, Suto S;
 DR WPI; 93-190118/24.
 PT Sexing bovine embryos - by amplifying specific DNA sequences by
 PT polymerase chain reaction using male-specific and gender-neutral
 PT primers
 PS Claim 1; Pages 9-10; 40pp; English.
 CC Q43719 is the DNA sequence of a plasmid which hybridises
 CC specifically to bovine male genomic (BMG) DNA. The plasmid having
 CC this DNA sequence was transfected into E. coli and the transfectant
 CC E.c.118-bms1 has been deposited as FERM BP-4095. Q43719 was used as
 CC a probe to screen a bovine male genomic library. The plaque
 CC hybridization gave 28 positive clones. One of them was picked up and
 CC its DNA was extracted and EcoRI digested into 2 fragments (Q43720,
 CC Q43721). Both DNAs specifically hybridize to the male DNA. They were
 CC introduced into E. coli and the resultant recombinants E.c.gem-bms1
 CC and E.c.gem-bms2 have been deposited as FERM BP-4089 and FERM BP-4090
 CC respectively. One clone was selected from the 28 clones (see
 CC Q43722). This DNA sequence hybridized to both male and female DNA,
 CC but images were different from each other. DNA having sequence
 CC Q43722 was introduced into E. coli and the recombinant E.c.gem-bms3
 CC has been deposited as FERM BP-4091. When male-specific clones were
 CC explored 20 clones were found to hybridize to both male and female
 CC DNA in the Southern blot analysis. Three clones which were expected
 CC to be repetitious were selected and used as probes for the Southern
 CC blot analyses (see Q43723, Q43724, Q43725). DNA consisting of these
 CC sequences were introduced into E. coli and the resultant
 CC recombinants, E.c.118-bmf1, E.c.118-bmf2, E.c.118-bmf3, have been
 CC deposited as FERM BP-4092, FERM BP-4093 and FERM BP-4094.
 SQ Sequence 1404 BP; 385 A; 289 C; 310 G; 420 T;

Query Match 70.0%; Score 14; DB 7; Length 1404;
 Best Local Similarity 93.8%; Pred. No. 2.28e+01;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 381 aacttcctaaaggga 396
 |||||
 QY 3 aacttcctaaaggga 18

RESULT 10
 ID Q98950 standard; DNA; 3134 BP.

AC Q98950;
 DT 25-APR-1996 (first entry)
 DE bg8 insert encoding a B-G subregion fragment of fowl MHC.
 KW Domesticated fowl; chicken; turkey; pheasants; B-G antigen; MHC;
 KW major histocompatibility complex; haplotyping; Marek's disease;
 KW restriction fragment length polymorphism; bg8 insert; probe; ds.
 OS Synthetic.
 PN US5451670-A.
 PD 19-SEP-1995.
 PF 30-JUN-1987; 068176.
 PR 30-JUN-1987; US-068176.
 PR 09-DEC-1987; US-130529.
 PR 23-JUN-1988; US-210405.
 PR 28-SEP-1989; US-413301.
 PR 27-SEP-1990; US-588922.
 PR 22-APR-1991; US-688326.
 PR 07-APR-1992; US-865662.
 PA (CITY) CITY OF HOPE.
 PI Miller MM;
 DR WPI; 95-336319/43.

PT New DNA encoding the B-G antigen of fowl MHC - and derived probes,
 PT useful for haplotyping and determining genotype at specific loci
 PS Claim 5; Columns 23-28; 61pp; English.
 CC The DNA sequences Q98944-55 contain a polymorphic portion of the
 CC coding region of domestic fowl's (i.e. chicken, turkey, pheasant)
 CC MHC B-G subregion (Q98944 encodes the B-G subregion fragment

```

RESULT 12
ID   T14694 standard; cDNA; 5342 BP.
AC   T14694;
DT   21-NOV-1996 (first entry)
DE   DNA polymerase.
KW   endonuclease 2; DNA polymerase; Pyrococcus furiosus; rare cutter;
KW   cleavage; cloning; sequencing; ds.
OS   Pyrococcus furiosus strain KOD1.
FH   Key
FH   Location/Qualifiers
FT   cds
FT   156..5168
FT   /*tag= a
FT   /product= DNA polymerase
FT   misc_feature
FT   1374..2453
FT   /*tag= b
FT   /note= "IVS-A"
FT   misc_feature
FT   2708..4316
FT   /*tag= c
FT   /note= "IVS-B; encodes endonuclease-2 (see T40731)"
FT   J08070864-A.
FT   19-MAR-1996.
FT   PD   19-MAR-1996.
FT   PF   02-SEP-1994; 209619.
FT   PR   02-SEP-1994; JP-209619.
FT   (TOYM ) TOYOBO KK.
FT   DR   WPI; 96-203145/21.
FT   DR   P-PSDB; R94616.
FT   PT   Endonuclease 2 derived from super-thermophilic archaeobacteria - has
FT   PT   rare cleavage site, for cutting DNA into large fragments
FT   PS   Disclosure; Page 7-14; 19pp; Japanese.
FT   CC   The present sequence encodes a DNA polymerase isolated from Pyrococcus
FT   CC   furiosus strain KOD1. Part of the sequence, designated IVS-B, encodes for
FT   CC   endonuclease 2 activity. The endonuclease 2 is an approx. 62 kDa protein,
FT   CC   and is a rare cutter (see T14693), cleaving DNA into large fragments.
FT   CC   Sequence 5342 BP; 1541 A; 1187 C; 1517 G; 1097 T;
FT   SQ

Query Match      70.0%; Score 14; DB 22; Length 5342;
Best Local Similarity 100.0%; Pred. No. 2.28e+01;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5219 ctcccttagggaa 5232
Cp 19 ctcccttagggaa 6
|||||

RESULT 13
ID   T28360 standard; cDNA; 5342 BP.
AC   T28360;
DT   19-NOV-1996 (first entry)
DE   DNA polymerase gene, contg. fragment encoding endonuclease I.
DE   KOD1 strain; super-thermophilic; archaeobacteria; DNA polymerase;
KW   endonuclease I; 3'-sticky end; genetic engineering; ds.
OS   Pyrococcus furiosus.
FH   Key
FH   Location/Qualifiers
FT   cds
FT   156..5168
FT   /*tag= a
FT   mat_peptide
FT   1374..2453
FT   /*tag= b
FT   /note= "claimed fragment encoding endonuclease I"
FT   J08070863-A.
FT   19-MAR-1996.
FT   PD   19-MAR-1996.
FT   PF   01-SEP-1994; 208631.
FT   PR   01-SEP-1994; JP-208631.
FT   (TOYM ) TOYOBO KK.
FT   DR   WPI; 96-203144/21.
FT   DR   P-PSDB; R97047.
FT   PT   Endonuclease I from super-thermophilic archaeobacteria, P.furiosus -
FT   PT   useful for genetic engineering
FT   PS   Claim 10; Pages 7-14; 18pp; Japanese.
FT   CC   The present sequence is the P.furiosus strain KOD1 (a
FT   CC   super-thermophilic archaeobacteria) DNA polymerase gene, which
FT   CC   contains a claimed fragment encoding endonuclease I. The
FT   CC   endonuclease has a mol. wt. of ca. 41 kD, cleaves double stranded
FT   CC   DNA into large fragments, leaves a 3'-sticky end and is therefore

```

CC useful in genetic engineering. Chromosomal DNA obtd. from a 95
 CC degrees C P. furiosus KOD1 culture was PCR amplified using primers
 CC designed and synthesised according to the base sequence of
 CC P. furiosus derived Pfu polymerase. The amplified fragment was used
 CC for Southern hybridisation against a restriction enzyme treated
 CC KOD1 chromosomal DNA, to give a DNA polymerase encoding fragment of
 CC ca. 4-7 kb.
 SQ Sequence 5342 BP; 1541 A; 1187 C; 1517 G; 1097 T;

Query Match 70.0%; Score 14; DB 22; Length 5342;
 Best Local Similarity 100.0%; Pred. No. 2.28e+01;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5219 ctcccttagggaa 5232
 Cp 19 ctcccttagggaa 6
 |||||

RESULT 14
 ID T46159 standard; DNA; 19932 BP.
 DE T46159;
 30-JUN-1997 (first entry)
 CagI locus.
 KW CagI; CagA; virulence factor; exporter molecule; homology; pti gene;
 KW Bordetella pertussis; VIR B4; Agrobacterium tumefaciens; invasion factor;
 KW Salmonella; type I strain; virulence; diagnosis; H. pylori; infection;
 KW vaccine; treatment; duodenal; gastric ulcer; active gastritis;
 KW adenocarcinoma; ss.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 FT 389..1579
 FT /tag= a
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds 1794..2168
 FT /tag= b
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds 3343..4491
 FT /tag= c
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds 4488..5426
 FT /tag= d
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (1778..2173)
 FT /tag= e
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (2591..3001)
 FT /tag= f
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (3020..3259)
 FT /tag= g
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (5508..6233)
 FT /tag= h
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (6218..6613)
 FT /tag= h
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (6562..7074)
 FT /tag= i
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (7007..7375)
 FT /tag= j
 FT /note= "putative open reading frame; no start codon

FT given"
 FT complement (7371..7802)
 FT /tag= k
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (7975..8454)
 FT /tag= l
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (8496..8918)
 FT /tag= m
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (8977..9762)
 FT /tag= n
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (9784..10575)
 FT /tag= o
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds 10557..12758
 FT /tag= p
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (12743..13420)
 FT /tag= q
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (13374..13742)
 FT /tag= r
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (13866..14129)
 FT /tag= s
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT cds complement (19296..19832)
 FT /tag= t
 FT /note= "putative open reading frame; no start codon
 FT given"
 FT W09633274-A1.
 PN 24-OCT-1996.
 PD 18-APR-1996; IB0343.
 PF 20-APR-1995; US-425194.
 PR 07-JUN-1995; US-477451.
 PA (BIOC-) BIOGINE SPA.
 PI Covacci A.
 DR WPI: 96-485780/48.
 DR P-PSDB; W06930-50.
 PT Helicobacter pylori CagI polynucleotide and related proteins - used
 PT in diagnosis and in vaccines for the treatment of H. pylori
 PT infection associated disease
 PS Claim 2; Fig 3A-R; 303pp; English.
 CC The present sequence is the complete nucleotide sequence of the CagI
 CC locus, including the true 5'-terminus of CagI. The sequence was
 CC constructed using overlapping clones (one of which also overlaps with
 CC the CagA region). The CagI region contains clusters of putative open
 CC reading frames (ORFs) with different polarities. The putative ORFs for
 CC this region are shown above. It is hypothesised that some of these ORFs
 CC may encode exporter molecules with homology to the pti genes of
 CC Bordetella pertussis and VIR B4 genes of Agrobacterium tumefaciens and
 CC for proteins with motifs shared by the purported invasion factors of
 CC Salmonella genus. The absence of the CagA gene in the type I strains is
 CC associated with the absence of CagI sequences (which may encode virulence
 CC factors restricted to type I strains). The CagI nucleotide sequence, its
 CC fragments and encoded proteins are used in the diagnosis of H. pylori
 CC (esp. H. pylori type I strain) infection in an individual and in vaccines
 CC (claimed) for the treatment of H. pylori infection associated with e.g.
 CC duodenal and gastric ulcers, severe forms of active gastritis (esp. type
 CC gastritis) and gastric adenocarcinoma.
 SQ Sequence 19932 BP; 6677 A; 3635 C; 3483 G; 6137 T;

Query Match 70.0%; Score 14; DB 29; Length 19932;
Best Local Similarity 100.0%; Pred. No. 2.28e+01;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5738 ccttagggaagtt 5751
Cp 16 ccttagggaagtt 3

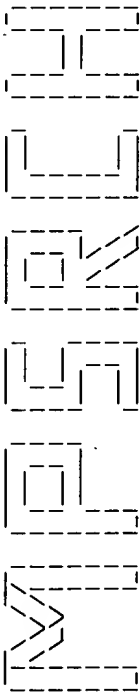
RESULT 15
ID V03751 standard; cDNA; 55 BP.
AC V03751;
DT 15-APR-1998 (first entry)
DE Porcine TNFalpha-converterase coding sequence fragment.
KW Tumour necrosis factor alpha convertase; TNFalpha-converterase; human; pig;
KW inhibitor; therapy; systemic inflammatory response syndrome; arthritis;
KW cardiovascular disease; infectious disease; inflammatory disease; AIDS;
KW porcine; restenosis; macular degeneration; multiple sclerosis; ss.
CC Sus scrofa.
CC W09735338-A2.
PF 02-OCT-1997.
PR 25-MAR-1997; E01497.
PA (GLAXO) GLAXO GROUP LTD.
PI Becherer JD, Chen W, Didsbury JR, Jin SC, McGeehan GM,
PI Moss ML, Rocque WJ, Schoenen FJ;
DR WPI: 97-489331/45.
PT Mammalian tumour necrosis factor alpha convertase - useful to screen
PT for new inhibitors which can treat disorders involving abnormal
PT levels of TNF alpha, e.g. inflammatory and cardiovascular disease
PS Example; Page 53; 132pp; English.
CC This sequence is a fragment of the porcine tumour necrosis factor
CC alpha (TNFalpha) convertase. The full length DNA is a coding sequence
CC of the invention, and was used to isolate the human TNFalpha-converterase
CC coding sequence (see V03745). TNFalpha-converterase can proteolytically
CC convert TNFalpha precursor to mature TNFalpha. The convertase can be used
CC to isolate novel compounds capable of binding to it, which preferably
CC inhibit its activity. Inhibitors of the convertase are useful to treat a
CC disease or condition characterised by an elevated level of TNFalpha in
CC the serum or tissues of a mammal, e.g. systemic inflammatory response
CC syndrome, reperfusion injury, cardiovascular disease, infectious disease,
CC obstetrical disorders, gynaecological disorders, inflammatory disease,
CC autoimmunity, allergic disease, atopic disease, malignancy, transplant
CC complication, septic shock, cachexia, AIDS, graft osteoporosis,
CC restenosis, psoriasis, infarction (preferably due to an ischaemic event),
CC rheumatoid arthritis, macular degeneration, osteoarthritis or multiple
CC sclerosis. The TNFalpha-converterase inhibitors can be modified for use as
CC ligands to purify TNFalpha-converterase.
CC Sequence 55 BP; 12 A; 11 C; 16 G; 16 T;

Query Match 65.0%; Score 13; DB 37; Length 55;
Best Local Similarity 93.3%; Pred. No. 8.10e+01;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 39 ctttcgggaagttcc 53
Cp 15 ctttcgggaagttcc 1

Search completed: Wed May 27 03:00:11 1998
Job time : 23 secs.

THIS PAGE BLANK (USPTO)



 (TM)

Release 3.0.5AA John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1997 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

March_on n.a. - n.a. database search, using Smith-Waterman algorithm
 On: Wed May 27 02:56:25 1998; MasPar time 58.38 Seconds
 Tabular output not generated. 571.704 Million cell updates/sec

Title: >SEQ1
 Description: (1-20) from new.seq
 Perfect Score: 20
 N.A. Sequence: 1 ccttgaaggatttcctcc 20
 Comp: ggaacttccttaagggaagg

Scoring table: TABLE default
 Gap 10

Nmatch STD : Dbase 0; Query 0
 Searched: 457423 seqs, 834342348 bases x 2
 Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb154
 1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
 7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro 12:em_v1
 Database: genbank106
 13:gb_ba 14:gb_htg 15:gb_in 16:gb_om 17:gb_ov 18:gb_pat
 19:gb_ph 20:gb_pl 21:gb_pr1 22:gb_pr2 23:gb_ro 24:gb_st
 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 6.658; Variance 3.016; scale 2.208

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	20	100.0	992	23	MMVCAM1B1	Mus musculus NIH Swiss	7.64e-02
2	20	100.0	2321	23	MUSVCAM01	Mus musculus vascular	7.64e-02
3	20	100.0	2355	22	S50587	VCAM1-vascular cell ad	7.64e-02
4	20	100.0	2396	21	HUMVCAMA	Human vascular cell ad	7.64e-02
5	20	100.0	2458	23	MMU42327	Mus musculus vascular	7.64e-02
6	20	100.0	5607	21	HUMVCAM1A	Human vascular cell ad	7.64e-02
7	16	80.0	145	23	MMCREB	Mus musculus CREB gene	2.74e+01
8	16	80.0	1007	23	MMU46027	Mus musculus CREB tran	2.74e+01
9	16	80.0	1023	23	RNCREB	Rat delta CREB mRNA fo	2.74e+01
10	16	80.0	1125	23	RNCREB	Rat mRNA for CAMP resp	2.74e+01
11	16	80.0	1161	23	MMTRANS	M.musculus mRNA for tr	2.74e+01
12	16	80.0	1175	13	LGU91638	Legionella gormanli ma	2.74e+01
13	16	80.0	1258	23	MUSCREB	Mus musculus CAMP resp	2.74e+01
14	16	80.0	4442	13	ATAGROSY	A.tumefaciens (C58) ac	2.74e+01
15	16	80.0	5414	25	AF049616	Cloning vector pFA2-CR	2.74e+01

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45		
c	16	16	16	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15		
	80.0	80.0	80.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0	75.0		
	33539	122228	146174	1549	1583	1662	2113	2319	2731	3254	4454	4648	4714	4839	4614	33085	33399	36503	40433	47334	64674	76169	80696	115967	117954	125974	186510	198456	210357			
	15	22	13	18	17	20	20	13	13	14	22	22	22	13	13	15	15	22	22	15	15	22	22	22	22	14	22	14	14	14	14	
	CELF54E7	AC000003	D90910	E01874	BSAMAB	FRGMTURF2	PSPLC	BSAMAG	S67784	AC003637	AB006756	AB006755	AB006757	BSORF4	AE000681	CELR02F2	CEC09H6	HUMU112E8	HSU22376	AC003054	AC004247	AC003014	HS384D21	AC002112	AC002112	HSAC002065	AC004253	HS451B15	AC003115	AC004166		
	Caenorhabditis elegans	Homo sapiens chromosom	Synechocystis sp. PCC6	DNA containing promote	B.stearothermophilus a	Rana catesbeiana mitoc	Haematococcus pluviali	Pisum sativum mRNA for	B.stearothermophilus a	N-carbamyl-L-amino aci	*** SEQUENCING IN PROG	Homo sapiens mRNA for	Homo sapiens mRNA for	Homo sapiens mRNA for	B.stearothermophilus o	Caenorhabditis elegans	Caenorhabditis elegans	Human cosmid U112E8, c	Human (c-myb) gene, co	Drosophila melanogaste	Drosophila melanogaste	Human PAC clone DJ290B	Human DNA sequence fro	Genomic sequence from	Human BAC clone RG021N	*** SEQUENCING IN PROG	Human DNA sequence fro	*** SEQUENCING IN PROG	*** SEQUENCING IN PROG			
	2.74e+01	2.74e+01	2.74e+01	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	1.07e+02	

ALIGNMENTS

RESULT LOCUS	1	MMVCAM1B1	992 bp	ROD	28-OCT-1995
DEFINITION		Mus musculus NIH Swiss vascular cell adhesion molecule-1 (VCAM-1)			
ACCESSION		U12878			
NID		g1041794			
KEYWORDS		1 of 6			
SEGMENT		mouse.			
SOURCE		Mus musculus			
ORGANISM		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.			
REFERENCE		1 (bases 1 to 796; 808 to 824)			
AUTHORS		Kumar,A.G., Dai,Y.X., Kozak,C.A., Mims,M.P., Gotto,A.M. and Ballantyne,C.M.			
TITLE		Murine VCAM-1: Molecular cloning, Mapping, and Analysis of a Truncated Form			
JOURNAL		Unpublished			
REFERENCE		2 (bases 1 to 992)			
AUTHORS		Kumar,A.G.			
TITLE		Direct Submission			
JOURNAL		Submitted (02-AUG-1994) Ajith G. Kumar, Department of Medicine, Baylor College of Medicine, 6535 Fannin Street, Houston, TX 77030, USA			
FEATURES		Location/Qualifiers			
Source		1..992			
		/organism="Mus musculus"			
		/strain="NIH Swiss"			
		/db_xref="taxon:10090"			
		/clone="lambda 1"			
		/clone.lib="Genomic library lambda FIX II, Stratagene, La Jolla, CA"			
		/chromosome="3"			
		/cell_line="NIH3T3"			
		/cell_type="fibroblast"			
		/dev_stage="adult"			
		632..636			
		TATA_signal			

sig_peptide 751..814
/gene="VCAM-1"
exon <751..814
/gene="VCAM-1"
/number=1
CDS 751..>814
/gene="VCAM-1"
/codon_start=1
/product="vascular cell adhesion molecule-1"
/db_xref="PID:g531863"
/translation="MPVKMVAVLGASTVILFAV"
BASE COUNT 280 a 204 c 204 g 304 t
ORIGIN

Query Match 100.0%; Score 20; DB 23; Length 992;
Best Local Similarity 100.0%; Pred. No. 7.64e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 592 CCTGAAGGATTTCCTCC 611
|||||
1 ccttgaaggatttcctcc 20

RESULT 2
LOCUS MUSVCAM01 2321 bp DNA ROD 23-MAR-1994

DEFINITION Mus musculus vascular cell adhesion molecule-1 (VCAM1) gene, exons 1 and 2.
ACCESSION L22301
NID g347974
KEYWORDS vascular cell adhesion molecule-1.
SEGMENT 1 of 8
SOURCE Mus musculus strain 129 DNA.
ORGANISM Mus musculus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2321)
AUTHORS Cybulsky,M.I., Allan-Motamed,M. and Collins,T.
TITLE Structure of the murine VCAM1 gene
JOURNAL Genomics 18 (2), 387-391 (1993)
MEDLINE 94117008
COMMENT MVI-6 individual isolate with similarities to MVI-9 and MVI-3; exon 1 encodes the signal peptide, and exon 2 encodes the first immunoglobulin-like domain of murine VCAM-1.
FEATURES
source
Location/Qualifiers
1..2321
/organism="Mus musculus"
/isolate="MVI-7"
/strain="129"
/sub_species="domesticus"
/isolate="MVI-9"
/isolate="MVI-6"
/db_xref="taxon:10090"
/cell_line="J1"
/cell_type="stem cells"
/dev_stage="blastula (blastocyst)"
/germline
/tissue_type="embryo"
/tissue_lib="lambda DASH II"
1..1340
/gene="VCAM1"
1317..1321
/gene="VCAM1"
1341..1434
/gene="VCAM1"
1341..1498
/gene="VCAM1"
/number=1
1499..1978
/gene="VCAM1"
/number=1
1979..2254

promoter
TATA_signal
5'UTR
exon
intron
exon

/gene="VCAM1"
/number=2
BASE COUNT 688 a 467 c 469 g 697 t
ORIGIN

Query Match 100.0%; Score 20; DB 23; Length 2321;
Best Local Similarity 100.0%; Pred. No. 7.64e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1277 CCTGAAGGATTTCCTCC 1296
|||||
1 ccttgaaggatttcctcc 20

RESULT 3
LOCUS S50587 2355 bp DNA PRI 22-FEB-1993

DEFINITION VCAM1=vascular cell adhesion molecule 1 [5' region, promoter] [human, Genomic, 2355 nt].
ACCESSION S50587
NID 9261284
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2355)
AUTHORS Neish,A.S., Williams,A.J., Palmer,H.J., Whitley,M.Z. and Collins,T.
TITLE Functional analysis of the human vascular cell adhesion molecule 1 promoter
JOURNAL J. Exp. Med. 176 (6), 1583-1593 (1992)
MEDLINE 93094762
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 120211] from the original journal article. This sequence comes from Fig. 1.
Map location: 1p31-32.

FEATURES
source
Location/Qualifiers
1..2355
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2355
/partial
/note="vascular cell adhesion molecule 1"
/gene="VCAM1"

BASE COUNT 728 a 447 c 431 g 746 t
ORIGIN

Query Match 100.0%; Score 20; DB 22; Length 2355;
Best Local Similarity 100.0%; Pred. No. 7.64e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2104 CCTGAAGGATTTCCTCC 2123
|||||
1 ccttgaaggatttcctcc 20

RESULT 4
LOCUS HUMVCAM 2396 bp DNA PRI 14-JAN-1995

DEFINITION Human vascular cell adhesion molecule-1 (VCAM1) gene, exon 1.
ACCESSION M92431
NID 9340197
KEYWORDS vascular cell adhesion molecule-1.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2396)
AUTHORS Iademarco,M.F., McQuillan,J.J., Rosen,G.D. and Dean,D.C.
TITLE Characterization of the promoter for vascular cell adhesion molecule-1 (VCAM-1)
JOURNAL J. Biol. Chem. 267 (23), 16323-16329 (1992)
MEDLINE 92355594


```

FEATURES
    source          Location/Qualifiers
    1..2396
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /map="lp32-p31"
    promoter
    1..2180
        /gene="VCAM1"
        /note="G00-127-922"
    gene
    1..2396
        /gene="VCAM1"
        /note="G00-127-922"
    TATA_signal
    2152..2157
        /gene="VCAM1"
        /note="G00-127-922"
    exon
    2300..2363
        /gene="VCAM1"
        /note="G00-127-922"
    intron
    2364..2396
        /gene="VCAM1"
        /note="G00-127-922"
    BASE COUNT      740 a 454 c 436 g 766 t
    ORIGIN
    Query Match      100.0%; Score 20; DB 21; Length 2396;
    Best Local Similarity 100.0%; Pred. No. 7.64e-02;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Db 2112 CCTGAAGGATTCCTCC 2131
    Qy 1 cctgaaggattccctcc 20
    |||||
    TATA_signal

    RESULT 5 MW42327 2458 bp DNA ROD 09-DEC-1997
    LOCUS Mus musculus vascular adhesion molecule-1 (VCAM-1) gene, partial
    cds.
    DEFINITION
    ACCESSION U42327
    NID g1147764
    KEYWORDS house mouse.
    ORGANISM Mus musculus
    Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
    Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    REFERENCE 1 (bases 1 to 2458)
    AUTHORS Korenaga,R., Ando,J., Kosaki,K., Isshiki,M., Takada,Y. and
    Kamiya,A.
    TITLE Negative transcriptional regulation of the VCAM-1 gene by fluid
    shear stress in murine endothelial cells
    JOURNAL Am. J. Physiol. 273 (5), C1506-C1515 (1997)
    MEDLINE 98042048
    REFERENCE 2 (bases 1 to 2458)
    AUTHORS Korenaga,R., Ando,J., Tsuboi,H. and Kamiya,A.
    TITLE Direct Submission
    JOURNAL Submitted (05-DEC-1995) Risa Korenaga, Cardiovascular Biomechanics,
    Faculty of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku,
    Tokyo 113, Japan
    Location/Qualifiers
    1..2458
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /cell_type="vascular endothelial cell"
    1896..1958
        /gene="VCAM-1"
    1896..1958
        /gene="VCAM-1"
    /codon_start=1
    /product="vascular adhesion molecule-1"
    /db_xref="PID:g1147765"
    /translation="MPVKMVAVLGASTVWLILFAV"
    BASE COUNT      718 a 482 c 482 g 776 t
    ORIGIN

```

```

Query Match      100.0%; Score 20; DB 23; Length 2458;
Best Local Similarity 100.0%; Pred. No. 7.64e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    Db 1738 CCTGAAGGATTCCTCC 1757
    Qy 1 cctgaaggattccctcc 20
    |||||
    RESULT 6 HUMVCAM1 5607 bp DNA PRI 14-JAN-1995
    LOCUS Human vascular cell adhesion molecule-1 (VCAM1) gene, complete cds.
    DEFINITION
    ACCESSION M73255
    NID g340195
    KEYWORDS immunoglobulin super gene family; leukocyte adhesion molecule;
    transmembrane protein; vascular cell adhesion molecule-1.
    SOURCE Homo sapiens (tissue library: EMBL3) peripheral blood DNA.
    ORGANISM Homo sapiens
    Eukaryotes; Mitochondrial eukaryotes; Metazoa; Chordata;
    Vertebrata; Eutheria; Primates; Catarrhini; Homidae; Homo.
    REFERENCE 1 (bases 1 to 5607)
    AUTHORS Cybulsky,M.I., Fries,J.W., Williams,A.J., Sultan,P., Eddy,R.,
    Byers,M., Shows,T., Gimbrone,M.A. Jr. and Collins,T.
    TITLE Gene structure, chromosomal location, and basis for alternative
    mRNA splicing of the human VCAM1 gene
    JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88 (17), 7859-7863 (1991)
    MEDLINE 91352090
    FEATURES
    Location/Qualifiers
    1..5607
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /tissue_type="peripheral blood"
        /tissue_lib="EMBL3"
        /map="lp32-p31"
        /gene="VCAM1"
        /note="G00-127-922"
        /evidence="experimental"
    join(519..706,779..1054,1127..1447,1520..1786,1859..2134,
    2207..2527,2897..3163,3236..3502,3922..4841)
    /gene="VCAM1"
    /note="G00-127-922"
    join(519..706,779..1054,1127..1447,1520..1786,1859..2134,
    2207..2527,2897..3163,3236..3502,3922..4841)
    /gene="VCAM1"
    519..706
        /gene="VCAM1"
        /note="G00-127-922"
        /number=1
    join(643..706,779..1054,1127..1447,1520..1786,1859..2134,
    2207..2527,2897..3163,3236..3502,3922..4082)
    /gene="VCAM1"
    /codon_start=1
    /number=2
    /db_xref="GDB:G00-127-922"
    /product="vascular cell adhesion molecule-1"
    /db_xref="PID:g340196"
    /translation="MPGKMVILGASNLIMFAASQAFKIETTPESRYLAQIGDSVS
    LTCSTTCGCESPFFSWRTQIDSPGLNGKVTNEGTTSTLTMPVSEFNEHSYLCATCESR
    KLEKIQVEIYSPFKDPEIHLGSLPAGKPIIVKCSVADVVPDRLEIDLKGDHLMK
    SQEFLDADRSLETSLEVTFTPVIEDIGKVLVCRAKLHIDEMSVPTVROAVKELQ
    VIISPKNTVISVNPSPKLBGGSVTWCSEGLPAPEIFWSKKLDNGLQHSNATL
    TLIAMRMEDSGIYEGVNLGNKRKEVELIVQEKFTVEISPGPRIAAGIGDSVMTL
    CSYNGCESPFSWRTQIDSPGLNGKVTNEGTTSTLTSPVSEFNEHSYLCATCESR
    EKGQVELYSPFKDPEIHLGSLPAGKPIIVKCSVADVVPDRLEIDLKGDHLMK
    EFLDADRSLETSLEVTFTPVIEDIGKVLVCRAKLHIDEMSVPTVROAVKELQ
    VIISPKNTVISVNPSPKLBGGSVTWCSEGLPAPEIFWSKKLDNGLQHSNATL
    VAPRDTVLVSPSSILLEGSSVNMTCISQGFAPKPIVMSRQLNGELQSEVATLTL
    ISYKMDSGVYLCEGNGQSRKEVELIIVTPKDKLTAFFSEKRVKGSQIRSLT
    CGNVPETWILKKKAEQGTVLKSIDGATYIRRAQLADAGVIECESKRVKGSQIRSLT
    LDYQGRNNKDYFPELLVLYFASLLIIPAIMIIVFARKANMKGSYSLVQAQSKV"
    779..1054
    exon

```

/gene="VCAM1"
 /note="G00-127-922"
 /number=2
 1127. .1447
 /gene="VCAM1"
 /note="G00-127-922"
 /number=3
 1520. .1786
 /gene="VCAM1"
 /note="G00-127-922"
 /number=4
 1859. .2134
 /gene="VCAM1"
 /note="G00-127-922"
 /number=5
 2207. .2527
 /gene="VCAM1"
 /note="G00-127-922"
 /number=6
 2897. .3163
 /gene="VCAM1"
 /note="G00-127-922"
 /number=7
 3236. .3502
 /gene="VCAM1"
 /note="G00-127-922"
 /number=8
 3922. .4841
 /gene="VCAM1"
 /note="G00-127-922"
 /number=9

BASE COUNT 1593 a 1081 c 1185 g 1726 t 22 others
 ORIGIN chromosome 1, map position lp31-32.

Query Match 100.0%; Score 20; DB 21; Length 5607;
 Best Local Similarity 100.0%; Pred. No. 7.64e-02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 455 CCTTGAAGGATTCCCTCC 474
 QY 1 ccttgaaggattccctcc 20
 |||||

RESULT 7 MCREBG 145 bp DNA ROD 14-JUL-1997
 LOCUS Mus musculus CREB gene for cAMP-responsive-element binding protein, exon 7.
 DEFINITION X67724 X65763
 ACCESSION 9288940
 ORDS CAMP responsive element binding protein; creb gene; sequence-specific binding protein.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 145)
 AUTHORS Ruppert,S.J.W.
 TITLE Direct Submission
 JOURNAL Submitted (04-AUG-1992) S.J.W. Ruppert, Howard Hughes Medical Institute, University of California at Berkeley, Dept of Molecular & Cell Biology, 401 Barker Hall, Berkeley, CA 94720, USA
 REFERENCE 2 (bases 1 to 145)
 AUTHORS Ruppert,S., Cole,T.J., Boshart,M., Schmid,E. and Schutz,G.
 TITLE Multiple mRNA isoforms of the transcription activator protein CREB: generation by alternative splicing and specific expression in primary spermatocytes
 JOURNAL EMBO J. 11 (4), 1503-1512 (1992)
 MEDLINE 92224889
 COMMENT Related sequences M95106, M95107 and X67718-X67728.
 FEATURES Location/Qualifiers
 source 1. .145

/organism="Mus musculus"
 /strain="c-ch/c-ch"
 /db_xref="taxon:10090"
 /cell_line="melanoma B16, teratocarcinoma F9"
 /map="2q32.3-q34"
 1. .145
 /gene="CREB"
 <1. .28
 /gene="CREB"
 /number=6
 /evidence=experimental
 29. .129
 /gene="CREB"
 /usedin=X67718:crebd_mRNA
 /usedin=X67718:creba_mRNA
 /usedin=X67718:crebg_mRNA
 /usedin=X67718:crebpsi_mRNA
 /usedin=X67719:crebd_cds
 /usedin=X67719:creba_cds
 /usedin=X67719:crebw_cds
 /number=7
 /evidence=experimental
 130. .>145
 /gene="CREB"
 /number=7
 /evidence=experimental
 47 a 23 c 31 g 44 t
 BASE COUNT
 ORIGIN

Query Match 80.0%; Score 16; DB 23; Length 145;
 Best Local Similarity 94.4%; Pred. No. 2.74e+01;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 98 AGGGAATCCTTCAAGG 115
 Cp 18 agggaaatccctccaagg 1
 |||||

RESULT 8 MMU46027 1007 bp mRNA ROD 25-JUL-1997
 LOCUS Mus musculus CREB transcription factor, novel spliced form, mRNA, partial cds.
 DEFINITION U46027
 ACCESSION 91655804
 NID
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1007)
 AUTHORS Yang,L., Lanier,E.R. and Kraig,E.
 TITLE Identification of a novel, spliced variant of CREB that is preferentially expressed in the thymus
 JOURNAL J. Immunol. 158 (6), 2522-2525 (1997)
 MEDLINE 97211815
 REFERENCE 2 (bases 1 to 1007)
 AUTHORS Kraig,E.
 TITLE Direct Submission
 JOURNAL Submitted (12-JAN-1996) Ellen Kraig, Univ. of Texas Health Science Center at San Antonio, Cellular & Structural Biology, 7703 Floyd Curl Dr., San Antonio, TX 78284, USA
 FEATURES Location/Qualifiers
 source 1. .1007
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="pmly2"
 /tissue_type="thymus"
 <1. .774
 /note="a novel CREB family member; CAMP response element-binding protein"

```

/codon_start=1
/function="transcription factor"
/product="CREB"
/db_xref="PID:g1655805"
/translation="EFETSRAVTEAENQMTVOAQPOIATLAQVMPAAHATSSAPTV
TLVLPNGRQMGVQQAQPSVQSPQVOTVOISIAESDSQESVDSVTSQKRR
EILSRYSIRKILNDLSDAPGVARIEEKESEETSPAITVTVPPIYOTSSGQT
ASGVQVQIOTIAPTSTIAPGVMASSPALPTQPAEAAARKREVRMLKREAAECRR
KKKEYVKLENRVLENQNKLTIELKALKDLYCHKSD"
BASE COUNT      316 a      244 c      215 g      232 t
ORIGIN

Query Match      80.0%; Score 16; DB 23; Length 1007;
Best Local Similarity 94.4%; Pred. No. 2.74e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 304 AGGGAATCCTTTCAAGG 321
|||||
18 agggaaatcccttcaagg 1

RESULT 9 RNCREB 1023 bp RNA ROD 07-AUG-1991
LOCUS

DEFINITION Rat delta CREB mRNA for CAMP-responsive element (CRE) binding
protein.
ACCESSION X60002
NID 956058
KEYWORDS CAMP response element; CAMP response element binding protein; CRE
binding protein; delta CREB gene.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1023)
AUTHORS Jungmann,R.A.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1991) R.A. Jungmann, Northwestern Univ Med School,
Dept of CMS Biology, 303 E. Chicago Avenue, Chicago, IL 60611, USA
2 (bases 1 to 1023)
REFERENCE Short,M.L., Manohar,C.F., Furtado,M.R., Ghadge,G.D., Wolinsky,S.M.,
Thimmappa,B. and Jungmann,R.A.
TITLE Nucleotide and derived amino-acid sequences of the CRE-binding
proteins from rat C6 glioma and HeLa cells
Nucleic Acids Res. 19 (15), 4290 (1991)
JOURNAL 9133414
EDLINE
COMMENT See also X60003, M27691, M34356 & X14788.
FEATURES
source
1..1023
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="C6 glioma cell ATCC CCL 107"
1..1023
/gene="delta CREB"
/evidence="experimental"
1..1023
/gene="delta CREB"
1..984
/gene="delta CREB"
/codon_start=1
/db_xref="PID:g56059"
/db_xref="SWISS-PROT:P15337"
/translation="MTMDSGADNQSGDAVTEAESQMTVOAQPOIATLAQVMPAA
HATSSAPTVTLVLPNGVQVHQVIAQAPSVIQSPQVOTVOISIAESDSQESVD
SVTSQKRRSPYKILNDLSDAPGVARIEEKESEETSPAITVTVPPIYOTSSGQT
YOTSSGQVIAITOGGAIQANNQDVGLOTLMTNAAATPGTTLIOYAOTTDGQO
ILVPSNOVVQAASDVQYQIARTAPTSTIAPGVMASSPALPTQPAEAAARKREVR
MKREAAECRRKKKEYVKLENRVLENQNKLTIELKALKDLYCHKSD"
346..360
/misc_feature
/gene="delta CREB"
/note="CAMP-dependent protein kinase consensus
phosphorylation site"

```

```

repeat_region 889..954
/note="leucine heptad repeat"
BASE COUNT      325 a      247 c      237 g      214 t
ORIGIN

Query Match      80.0%; Score 16; DB 23; Length 1023;
Best Local Similarity 94.4%; Pred. No. 2.74e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 331 AGGGAATCCTTTCAAGG 348
|||||
18 agggaaatcccttcaagg 1

Cp 18 agggaaatcccttcaagg 1

RESULT 10 RNCREB 1125 bp RNA ROD 12-SEP-1993
LOCUS

DEFINITION Rat mRNA for CAMP response element binding protein (CREB).
ACCESSION X14788
NID 956004
KEYWORDS CAMP response element binding protein; DNA binding protein; nuclear
protein; transcription factor.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1125)
AUTHORS Gonzalez,G.A., Yamamoto,K.K., Fischer,W.H., Karr,D., Menzel,P.,
Biggs,W. III., Vale,W.W. and Montminy,M.R.
TITLE A cluster of phosphorylation sites on the cyclic AMP-regulated
nuclear factor CREB predicted by its sequence
Nature 337 (6209), 749-752 (1989)
JOURNAL 89143746
EDLINE
COMMENT Data kindly reviewed (12-APR-1990) by Montminy M.R.
FEATURES
source
1..1125
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cell_line="PC12"
/codon_start=1
/codon_start=1
/codon_start=1
/db_xref="PID:g56005"
/db_xref="SWISS-PROT:P15337"
/translation="MTMDSGADNQSGDAVTEAESQMTVOAQPOIATLAQVMPAA
HATSSAPTVTLVLPNGVQVHQVIAQAPSVIQSPQVOTVOISIAESDSQESVD
SVTSQKRRSPYKILNDLSDAPGVARIEEKESEETSPAITVTVPPIYOTSSGQT
YOTSSGQVIAITOGGAIQANNQDVGLOTLMTNAAATPGTTLIOYAOTTDGQO
ILVPSNOVVQAASDVQYQIARTAPTSTIAPGVMASSPALPTQPAEAAARKREVR
MKREAAECRRKKKEYVKLENRVLENQNKLTIELKALKDLYCHKSD"
354 a      270 c      264 g      237 t
BASE COUNT      354 a      270 c      264 g      237 t
ORIGIN

Query Match      80.0%; Score 16; DB 23; Length 1125;
Best Local Similarity 94.4%; Pred. No. 2.74e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 412 AGGGAATCCTTTCAAGG 429
|||||
18 agggaaatcccttcaagg 1

Cp 18 agggaaatcccttcaagg 1

RESULT 11 MMTRANS 1161 bp RNA ROD 04-JUN-1996
LOCUS

DEFINITION M.musculus mRNA for transcription factor/DNA binding protein.
ACCESSION X92497
NID 91134858
KEYWORDS CREB protein; DNA binding protein; transcription factor.
SOURCE house mouse.

```

```

ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1161)
AUTHORS Blandy,J.A., Kaestner,K.H., Schmid,W., Gass,P. and Schutz,G.
TITLE Targeting of the CREB gene leads to up-regulation of a novel CREB
mRNA isoform
JOURNAL EMBO J. 15 (5), 1098-1106 (1996)
MEDLINE 96183194
REFERENCE 2 (bases 1 to 1161)
AUTHORS Kaestner,K.H.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1995) K.H. Kaestner, DKFZ, German Cancer Research
Center, Molecular Biology of the Cell I, Im Neuenheimer Feld 280,
69120 Heidelberg, FRG
FEATURES
source
1. .1161
/organism="Mus musculus"
/strain="12901a"
/db_xref="taxon:10090"
/cell_type="brain"
7. .>1161
201. .1064
/function="transcriptional factor/DNA binding protein"
/notes="unnamed protein product"
/codon_start=1
/db_xref="PID:e206536"
/db_xref="PID:g1134859"
/translation="MPAAHATSSAPTVTLVQLPNGOTVGVHGVIAOAPSIVQSPQVQ
TVOISTAESDSQESVDTSOKREILSRPSYRKTLNDLSSDAPGVPRIEREKS
EEETSAPATTVTPPIQVTSGGQIATQSGAIQANGIDGVQGLTFLWNA
TQPTGTLTAQTDCCQLLVPSNVVQAAGDVQIQTAPSTIAPGVVMASSP
ALPTQPAEAAARKREVLKMKNEARECRKKKEYVKLENRVAVLENGKTLIEELK
ALKDLXCHKSD"
BASE COUNT 333 a 282 c 302 g 244 t
ORIGIN
Query Match 80.0%; Score 16; DB 23; Length 1161;
Best Local Similarity 94.4%; Pred. No. 2.74e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 411 AGGGAATCTTTCACGS 428
Cp 18 agggaaatccctcaagg 1
RESULTS
DEFINITION Legionella gormanii macrophage infectivity potentiator (mip) gene,
complete cds.
ACCESSION U91638
NID 9231693
KEYWORDS Legionella gormanii.
SOURCE Legionella gormanii.
ORGANISM Eubacteria; Proteobacteria; gamma subdivision; Legionellaceae;
Legionella.
REFERENCE 1 (bases 1 to 1175)
AUTHORS Ratcliff,R.M., Donnellan,S.C., Lanser,J.A., Manning,P.A. and
Heuzenroeder,M.W.
TITLE Interspecies sequence differences in the Mip protein from the genus
Legionella: implications for function and evolutionary relatedness
Mol. Microbiol. 25 (6), 1149-1158 (1997)
JOURNAL 98010353
MEDLINE
REFERENCE 2 (bases 1 to 1175)
AUTHORS Ratcliff,R.M., Donnellan,S.C., Lanser,J.A., Manning,P.A. and
Heuzenroeder,M.W.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-1997) Infectious Diseases Laboratories, Institute
of Medical and Veterinary Science, Frome Road, Adelaide, South
Australia 5000, Australia

```

```

FEATURES
source
1. .1175
/organism="Legionella gormanii"
/strain="ATCC 33297"
/db_xref="taxon:40334"
335. .1033
/gene="mip"
335. .1033
/gene="mip"
/notes="peptidyl-prolyl cis/trans isomerase; PPIase; FKBP
immunophilin"
/codon_start=1
/product="macrophage infectivity potentiator"
/db_xref="PID:g2231694"
/translation="MKMLVTAAMINGLAMSTAMATATSLVTDKDKLSYSTGADLGK
NFKNQIDINPEALAKGQDQSGAQLILTEQMKDVLNKFQDKLMKRSABFNKAE
ENKSGEAPFLSTNKSQGVVLPVSLQYKVTGAGTSGKPGKSDTVTVVSTGLDGTV
FDSTEKTKPATFQVSVIPGWTEALQLMPAGSTWEIIVPSDLAYGPRSVGGPIGPNE
TLIFKHLISVYKKA"
BASE COUNT 380 a 180 c 235 g 380 t
ORIGIN
Query Match 80.0%; Score 16; DB 13; Length 1175;
Best Local Similarity 100.0%; Pred. No. 2.74e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1041 GAAGGATTTCCTCC 1056
QY 5 gaaggatttcctcc 20
RESULTS
DEFINITION Mus musculus CAMP response element binding protein (CREB1) mRNA,
complete cds.
ACCESSION M95106
NID 9192713
KEYWORDS CAMP responsive element binding protein.
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ruppert,S., Cole,T.J., Boshart,M., Schmid,E. and Schutz,G.
TITLE Multiple mRNA isoforms of the transcription activator protein CREB:
Generation by alternative splicing and specific expression in
primary spermatocytes
EMBO J. 11, 1503-1512 (1992)
JOURNAL 92224889
MEDLINE
REFERENCE 2 (bases 1 to 1258)
AUTHORS Cole,T.J., Copeland,N.G., Gilbert,D.J., Jenkins,N.A., Schuetz,G.
and Ruppert,S.
TITLE The mouse CREB (CAMP responsive element binding protein) gene:
Structure, promoter analysis, and chromosomal localization
Genomics 13, 974-982 (1992)
JOURNAL 92372067
MEDLINE
FEATURES
source
1. .1258
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="B16 melanoma/F9 teratocarcinoma"
/map="1"
1. .135
/gene="CREB1"
1. .1258
/gene="CREB1"
136. .1161
/gene="CREB1"
/codon_start=1
/product="CAMP response element binding protein"
5'UTR
gene
CDS

```


Query Match 80.0%; Score 16; DB 25; Length 5414;
Best Local Similarity 94.4%; Pred. No. 2.74e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1467 AGGGAATCCTTTCAAGG 1484
|||||||
Cp 18 agggaaatcccttcaagg 1

Search completed: Wed May 27 02:57:31 1998
Job time : 66 secs.

 M E S R L H
 (TM)

Release 3.0.5AA John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1997 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

mpsrch_nm n.a. - n.a. database search, using Smith-Waterman algorithm
 On: Wed May 27 02:57:49 1998; Maspar time 17.23 Seconds
 147.868 Million cell updates/sec
 Tabular output not generated.

Title: >SEQ1
 Description: (1-20) from new.seq
 Perfect Score: 20
 N.A. Sequence: 1 ccttgaagggtttccctcc 20
 Comp: ggaacttccttaaggagg

Scoring table: TABLE default
 Gap 10

Nmatch STD : Dbase 0; Query 0
 Searched: 176923 seqs, 63680241 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-geneseq31-2
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37

Statistics: Mean 5.229; Variance 2.864; scale 1.826

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	20	100.0	20	30	T69675 Transcription factor	1.08e-02
2	20	100.0	36	15	Q90101 VCAM-1 expression inh	1.08e-02
3	20	100.0	47	15	Q90100 VCAM-1 expression inh	1.08e-02
4	20	100.0	1032	1	Q06690 5' UTR of VCAM1 from	1.08e-02
5	18	90.0	19	15	Q90103 VCAM-1 expression inh	1.75e-01
6	16	80.0	91	9	Q51746 Oligonucleotide probe	2.56e+00
7	16	80.0	91	9	Q51746 Oligonucleotide probe	2.56e+00
8	15	75.0	34	15	Q90105 VCAM-1 expression inh	9.39e+00
9	15	75.0	378	29	T67638 H. pylori flagella-as	9.39e+00
10	15	75.0	387	29	T68240 H. pylori flagella-as	9.39e+00
11	15	75.0	1230	25	T39170 Bacillus stearothermo	9.39e+00
12	15	75.0	1548	1	N80038 Colony stimulating ge	9.39e+00
13	15	75.0	1549	1	N80398 Region upstream of co	9.39e+00
14	15	75.0	1629	21	T18011 Beta-ionone 4-methyle	9.39e+00

15	15	75.0	1860	4	Q36728	Heat resistant carbam	9.39e+00
16	14	70.0	34	15	O90105	VCAM-1 expression inh	3.33e+01
17	14	70.0	238	20	T22901	Human gene signature	3.33e+01
18	14	70.0	316	20	T22912	Human gene signature	3.33e+01
19	14	70.0	534	18	T19067	Human gene signature	3.33e+01
20	14	70.0	1438	32	T79126	Human serine protease	3.33e+01
21	14	70.0	2352	14	Q42475	Truncated FLT (sVEGF-	3.33e+01
22	14	70.0	2427	12	Q70435	Human glyco-protein 1	3.33e+01
23	14	70.0	2475	1	Q04780	Sequence encoding CAM	3.33e+01
24	14	70.0	2523	33	T62101	cDNA encoding amino-t	3.33e+01
25	14	70.0	2523	36	V01457	Human VEGF receptor e	3.33e+01
26	14	70.0	2651	14	Q42468	SVGEF-RI gene.	3.33e+01
27	14	70.0	4458	34	T68840	Photorehabdus luminesc	3.33e+01
28	14	70.0	8357	34	T91532	Joining region betwee	3.33e+01
29	13	65.0	21	6	Q37156	Probe to detect G-CSF	1.13e+02
30	13	65.0	36	10	Q45406	Oligonucleotide formi	1.13e+02
31	13	65.0	253	21	T24716	Human gene signature	1.13e+02
32	13	65.0	308	2	N70221	Plasmid pKCS-12 inser	1.13e+02
33	13	65.0	525	1	Q04481	Plasmid pAS28 encodin	1.13e+02
34	13	65.0	525	34	T64619	G-CSF receptor agonis	1.13e+02
35	13	65.0	534	34	T64609	G-CSF receptor agonis	1.13e+02
36	13	65.0	534	34	T64610	G-CSF receptor agonis	1.13e+02
37	13	65.0	797	11	O89201	Human Pancreatititis-As	1.13e+02
38	13	65.0	849	29	T67890	H. pylori flagella as	1.13e+02
39	13	65.0	1083	29	T41787	Fusion peptide #2 hav	1.13e+02
40	13	65.0	1315	1	N90022	Human granulocyte col	1.13e+02
41	13	65.0	1530	2	N70222	Plasmid pBRG4 insert.	1.13e+02
42	13	65.0	2455	7	Q45988	G-CSF-(Gly)4-HSA chim	1.13e+02
43	13	65.0	2960	2	N70224	Gene encoding human G	1.13e+02
44	13	65.0	3069	2	Q11772	Sequence encoding hum	1.13e+02
45	13	65.0	4268	5	Q31000	Notch clone hN5k full	1.13e+02

ALIGNMENTS

RESULT 1
 ID T69675 standard; DNA; 20 BP.
 AC T69675;
 DT 04-AUG-1997 (first entry)
 DE Transcription factor NF-kappa-B DNA binding site antagonist.
 KW Decoy; antagonist; NF-kappa-B; NF-kB; transcription; regulation;
 KW prevention; treatment; disease; ischaemia; ischaemia; inflammation;
 KW autoimmune; cancer; metastasis; cachexia; organ; transplantation;
 KW surgery; ds.
 OS Synthetic.
 PN WO9635430-A1.
 PD 14-NOV-1996.
 PF 10-MAY-1996; J01234.
 PR 12-MAY-1995; JP-114990.
 PR 02-NOV-1995; JP-285504.
 PA (FUJI) FUJISAWA PHARM CO LTD.
 PI Chiba T, Kawamura I, Maeda K, Morishita R, Ogiwara T;
 PI Sugimoto T;
 DR WPI; 96-518400/31.
 PT Anti-sense NF-kB agent - for treatment of ischaemia, inflammatory
 PT disease auto-immune disease, etc.
 PS Claim 7; Page 9; 18pp; Japanese.
 CC The present sequence is a decoy, which specifically antagonises
 CC the nucleic acid site to which a NF-kappa-B transcription regulator
 CC binds. It can be used to prevent or treat diseases caused by
 CC NF-kappa-B, e.g. ischaemia, inflammatory and autoimmune disease,
 CC cancer metastasis and cachexia, especially following organ
 CC transplant or surgery.
 SQ Sequence 20 BP; 3 A; 7 C; 4 G; 6 T;

Query Match 100.0%; Score 20; DB 30; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.08e-02;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ccttgaagggtttccctcc 20
 ||||||||||||||||
 Qy 1 ccttgaagggtttccctcc 20

```

RESULT 2
ID Q90101 standard; DNA; 36 BP.
AC Q90101;
DE 11-JAN-1996 (first entry)
DE VCAM-1 expression inhibiting oligonucleotide.
DE Vascular cell adhesion molecule; VCAM-1; inhibitory oligonucleotide;
KW transcriptional regulatory factor; diagnosis; treatment; restenosis;
KW atherosclerosis; inflammatory disease; ds.
OS Synthetic.
PN WO9512415-A1.
PD 11-MAY-1995.
PF 07-NOV-1994; U12797.
PR 05-NOV-1993; US-147878.
PA (ISIS-) ISIS PHARM INC.
PA (UEM-) UNIV EMORY.
PI Bennett CF, Medford RM;
PI WPI; 95-193802/25.
DR Oligo-nucleotide(s) which modulate vascular cell adhesion molecule
PT expression by binding a transcriptional regulatory element - used to
PT diagnose and treat atherosclerosis, restenosis or inflammatory
PT disease
PT Claim 17; Page 33; 49pp; English.
CC Q90100-Q90111 bind the vascular cell adhesion molecule (VCAM)-1
CC gene transcriptional regulatory factor, therefore inhibiting the
CC expression of VCAM-1. They can be used in the diagnosis and
CC treatment of restenosis, atherosclerosis and inflammatory
CC diseases.
SQ Sequence 36 BP; 3 A; 13 C; 9 G; 11 T;

Query Match 100.0%; Score 20; DB 15; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.08e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 16 ccttgaaggattccctcc 35
|||||
QY 1 ccttgaaggattccctcc 20

RESULT 3
ID Q90100 standard; DNA; 47 BP.
AC Q90100;
DE 11-JAN-1996 (first entry)
DE VCAM-1 expression inhibiting oligonucleotide.
DE Vascular cell adhesion molecule; VCAM-1; inhibitory oligonucleotide;
KW transcriptional regulatory factor; diagnosis; treatment; restenosis;
KW atherosclerosis; inflammatory disease; ds.
OS Synthetic.
PN WO9512415-A1.
PD 11-MAY-1995.
PF 07-NOV-1994; U12797.
PR 05-NOV-1993; US-147878.
PA (ISIS-) ISIS PHARM INC.
PA (UEM-) UNIV EMORY.
PI Bennett CF, Medford RM;
PI WPI; 95-193802/25.
DR Oligo-nucleotide(s) which modulate vascular cell adhesion molecule
PT expression by binding a transcriptional regulatory element - used to
PT diagnose and treat atherosclerosis, restenosis or inflammatory
PT disease
PT Claim 17; Page 33; 49pp; English.
CC Q90100-Q90111 bind the vascular cell adhesion molecule (VCAM)-1
CC gene transcriptional regulatory factor, therefore inhibiting the
CC expression of VCAM-1. They can be used in the diagnosis and
CC treatment of restenosis, atherosclerosis and inflammatory
CC diseases.
SQ Sequence 47 BP; 3 A; 17 C; 12 G; 15 T;

Query Match 100.0%; Score 20; DB 15; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.08e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 22 ccttgaaggattccctcc 41

```

```

|||||
QY 1 ccttgaaggattccctcc 20

RESULT 4
ID Q06690 standard; DNA; 1032 BP.
AC Q06690;
DE 27-FEB-1991 (first entry)
DE 5' UTR of VCAM1 from clone VCL-16.
DE Endothelial cell-leucocyte adhesion molecule 1; ELAM1;
KW vascular cell adhesion molecule 1; VCAM1; UTR; vector; promoter;
KW molecule involved in leucocyte adhesion; MILA; ss.
FH Key Location/Qualifiers
FT tata_signal 261..266
FT /tag= a
FT protein_bind 213..222
FT /tag= b
FT /number= 1
FT /note="NF-kappaB binding sequence"
FT protein_bind 227..236
FT /tag= c
FT /number= 2
FT /note="NF-kappaB binding sequence"
FT exon 303..472
FT /tag= d
FT /number= 2
FT /note="cDNA clone 41"
FT intron 473..1023
FT /tag= e
FT /number= 1
FT exon 1024..1032
FT /tag= f
FT /number= 2
PN WO9013300-A.
PD 15-NOV-1990.
PF 27-APR-1990; U02357.
PR 28-APR-1989; US-345151.
PR 01-JUN-1989; US-359516.
PR 18-DEC-1989; US-452675.
PA (BIOG-) BIOGEN INC.
PI Hession C, Lobb RR, Goelz SE, Born L, Benjamin CD;
PI Rosa MD;
PI WPI; 90-361248/48.
DR Endothelial cell adhesion mols. - MILAs and DNA encoding them and
PT inhibition-detection of binding of leukocytes to endothelial
PT cells
PS Disclosure; Fig 8; 136pp; English.
CC The 9 nucleotides N between bases 4013 and 4024 were not legible in
CC the specification.
CC VCAM1 clones were isolated by screening a human genomic EMBL3
CC library with a 32P-labeled 30 base oligomer probe homologous to the
CC 5' end of the VCAM1 cDNA.
CC NF-kappaB DNA binding activity is stimulated in endothelial cells
CC by IL-1 and TNF.
CC The promoter sequence is useful to construct vectors inducible by
CC cytokines, and bacterial LPS, or other agents found to induce
CC expression of ELAMs in endothelial cells. Such vectors may be useful
CC in gene transfer assays wherein the inducible promoter is positioned
CC so that it drives transcription of a reporter gene such as
CC chloramphenicol.
CC See also Q06686-91.
SQ Sequence 1032 BP; 273 A; 192 C; 216 G; 342 T;

Query Match 100.0%; Score 20; DB 1; Length 1032;
Best Local Similarity 100.0%; Pred. No. 1.08e-02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 221 ccttgaaggattccctcc 240
|||||
QY 1 ccttgaaggattccctcc 20

RESULT 5

```



```

ID T67638 standard; DNA; 378 BP.
AC T67638;
DT 11-JUL-1997 (first entry)
DE H. pylori flagella-associated protein ORF 3942217.aa.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW Identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT cds 1..291
FT /tag= a
FT /transl_except= (pos: 223..225, aa: Xaa)
FT /note= "Xaa - unknown, no stop codon given"
PN W09640893-Al.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR ) ASTRA AB.
PA Berglindh OT, Smith D, Mellgaard BL;
P-PSDB; W20473.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 1; Page 278; 1481pp; English.
CC This sequence encodes a H. pylori flagella-associated protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 378 BP; 121 A; 59 C; 94 G; 103 T;

Query Match 75.0%; Score 15; DB 29; Length 378;
Best Local Similarity 94.1%; Pred. No. 9.39e+00;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 350 ggagggaatccattca 366
| | | | | | | | | | | | | | | |
20 ggagggaatccattca 4

RESULT 10
ID T68240 standard; DNA; 387 BP.
AC T68240;
DT 21-JUL-1997 (first entry)
DE H. pylori flagella-associated protein ORF hpell122orf5.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW Identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis; ds.
OS Helicobacter pylori.
FH Key Location/Qualifiers
FT cds 1..387
FT /tag= a
FT /note= "no stop codon given"
PN W09640893-Al.
PD 19-DEC-1996.
PF 06-JUN-1996; U09122.
PR 07-JUN-1995; US-487032.
PR 01-APR-1996; US-630405.
PA (ASTR ) ASTRA AB.
PA Berglindh OT, Smith D, Mellgaard BL;

```

```

DR WPI; 97-052306/05.
DR P-PSDB; W20987.
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
PS Claim 1; Page 984; 1481pp; English.
CC The present sequence encodes a H. pylori flagella-associated protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
SQ Sequence 387 BP; 126 A; 60 C; 96 G; 105 T;

Query Match 75.0%; Score 15; DB 29; Length 387;
Best Local Similarity 94.1%; Pred. No. 9.39e+00;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 359 ggagggaatccattca 375
| | | | | | | | | | | | | | | |
20 ggagggaatccattca 4

RESULT 11
ID T39170 standard; DNA; 1230 BP.
AC T39170;
DT 04-MAR-1997 (first entry)
DE Bacillus stearothermophilus amino acid amidohydrolase DNA.
KW Amino acid amidohydrolase; carbamoylase; L-methionine; stereospecific;
KW thermostable; N-carbamoyl; cam gene; ds.
OS Bacillus stearothermophilus.
FH Key Location/Qualifiers
FT cds 1..1230
FT /tag= a
FT /product= amino_acid_amidohydrolase
FT /transl_except= pos:1..3, aa:Met
FT /note= "TTG initiation codon"
PN FR2728905-Al.
PD 05-JUL-1996.
PF 29-DEC-1994; 015838.
PR 29-DEC-1994; FR-015838.
PA (RHON ) RHONE POULENC NUTRITION ANIMALE.
PI Batisse N, Dion M, Hallet JN, Lecocq FM, Sakanyan V;
PI Weigel P;
DR WPI; 96-335881/34.
DR P-PSDB; W03544.
PT New stereospecific, heat-stable amino acid amidohydrolase - from
PT B.stearothermophilus and related DNA, esp. for prodn. of L-Met.
PS Claim 4; Page 9-11; 17pp; French.
CC PSTI fragments of Bacillus stearothermophilus NCIB 8224 genomic DNA
CC were inserted into pBR322 and used for transforming E.coli hosts.
CC Selection was on medium which contained tetracycline but lacked
CC arginine. All positive transformants carried plasmids with a 4.7 kb
CC insert. The present sequence is a fragment of the insert. Amino
CC acid amidohydrolase encoded by the fragment is stereospecific and
CC heat-stable (maximum activity at 55-60 deg.C). The enzyme
CC hydrolyses N-carbamoyl amino acid derivs. to L-amino acids and is
CC particularly useful for production of L-methionine.
SQ Sequence 1230 BP; 268 A; 312 C; 413 G; 237 T;

Query Match 75.0%; Score 15; DB 25; Length 1230;
Best Local Similarity 100.0%; Pred. No. 9.39e+00;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 460 gaagggaattccctc 474

```

```

Qy      5 gaaggattccctc 19
|||||
RESULT 12
ID      N80038 standard; cDNA; 1548 BP.
AC      N80038;
DT      30-OCT-1990 (first entry)
DE      Colony stimulating gene region and promoter region.
KW      Colony stimulating factor; promoter.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      misc_feature 796..1315
FT      /tag= a
FT      /note="claimed CSF-1 gene promoter region"
PN      DE3808213-A.
PD      22-SEP-1988.
PF      11-MAY-1988; 808213.
PR      13-MAR-1987; JP-056890.
PA      (GREC) Green Cross Corp.
PI      Murakami K, Nakakubo H, Kaneda T, Nagai M, Arimura H;
DR      WPI; 88-272091/39.
PT      New promoter sequence of colony stimulating factor gene -
PS      for controlling expression of protein in mammalian cells.
CC      Disclosure; d claim 1, page 2; 12pp; german.
CC      The colony stimulating factor-1 gene promoter region is useful for
CC      controlling expression of urokinase, hepatitis B antigen, human serum
CC      albumin and interferons in, eg mouse L cells or cell lines derived from
CC      T cells and tumours.
SQ      Sequence 1548 BP; 349 A; 388 C; 498 G; 313 T;

Query Match 75.0%; Score 15; DB 1; Length 1548;
Best Local Similarity 89.5%; Pred. No. 9.39e+00;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db      862 cttcaaggattccctcc 880
|||||
Qy      2 cttgaaggattccctcc 20

RESULT 13
ID      N80398 standard; DNA; 1549 BP.
AC      N80398;
DT      11-NOV-1990 (first entry)
DE      Region upstream of colony stimulating factor-1 gene.
KW      Colony stimulating factor-1; promoter region.
OS      Homo sapiens.
FH      Key      Location/Qualifiers
FT      misc_feature 797..1316
FT      /tag= a
FT      /note="claimed promoter sequence"
FT      misc_feature 1317..
FT      /tag= b
FT      /note="known sequence"
FT      misc_feature 1193..1240
FT      /tag= c
FT      /note="G-T cluster"
FT      misc_feature 1063..1071
FT      /tag= d
FT      /note="region homologous to enhancer-core region of
FT      immunoglobulin heavy chain gene"
FT      misc_feature 1186..1194
FT      /tag= e
FT      /note="region homologous to enhancer-core region of
FT      immunoglobulin heavy chain gene"
FT      misc_feature 797..806
FT      /tag= f
FT      /note="region homologous to consensus sequence upstream
FT      from lymphokine gene"
FT      misc_feature 869..878
FT      /tag= g
FT      /note="region homologous to consensus sequence upstream
FT      from lymphokine gene"
FT      /tag= h
FT      /note="region homologous to consensus sequence upstream
FT      from lymphokine gene"
FT      misc_feature 1495..
FT      /tag= i
FT      1534..
FT      /tag= j
FT      intron
FT      GB2204042-A.
FT      02-NOV-1988.
PD      11-MAR-1988; 005799.
PR      13-MAR-1987; JP-056890.
PA      (GREC) Green Cross Corp.
PI      Murakami K, Nakakubo H, Kaneda T, Nagai M, Arimura H;
DR      WPI; 88-272091/39.
PT      New promoter sequence of colony stimulating factor gene -
PS      for controlling expression of protein in mammalian cells.
CC      Disclosure; pp; English.
CC      The DNA contains a claimed CSF-1 promoter region, which is useful for
CC      controlling expression of urokinase, hepatitis B antigen, human serum
CC      albumin and interferons, etc. in mouse L cells or cell lines derived from
CC      T cells and tumours.
SQ      Sequence 1549 BP; 349 A; 388 C; 499 G; 313 T;

Query Match 75.0%; Score 15; DB 1; Length 1549;
Best Local Similarity 89.5%; Pred. No. 9.39e+00;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db      863 cttcaaggattccctcc 881
|||||
Qy      2 cttgaaggattccctcc 20

RESULT 14
ID      T18011 standard; cDNA; 1629 BP.
AC      T18011;
DT      03-OCT-1996 (first entry)
DE      Beta-ionone 4-methylene gp. to keto gp. converting enzyme cDNA.
KW      Converting enzyme; 4-methylene group; beta-ionone compound;
KW      keto group; microbial host cell; transformation; recombinant;
KW      vector; production; astaxanthin; 4-ketozeaxanthin; canthaxanthin;
KW      echinenone; ketocarotenoids; NIES-144; biosynthesis; ds.
OS      Haematococcus pluvialis.
FH      Key      Location/Qualifiers
FT      cds 168..1082
FT      /tag= a
FT      WO9606172-A1.
FT      29-FEB-1996.
PD      18-AUG-1995; J01640.
PR      23-AUG-1994; JP-198775.
PR      19-SEP-1994; JP-223798.
PR      07-MAR-1995; JP-047266.
PA      (KIRI) KIRIN BEER KK.
PI      Kajiwara S, Kondo K, Misawa N;
DR      WPI; 96-151375/15.
DR      P-PSDB: R92096.
PT      Polypeptide(s) and their DNA which introduce a keto-gp. in
FT      beta-ionone cpds. - for prodn. of keto:carotenoid(s) in
FT      transformed E. coli.
PS      Claim 10; Pages 33-36; 63pp; Japanese.
CC      The present sequence encodes a converting enzyme, which converts
CC      the 4-methylene gp. of a beta-ionone cpd. into a keto gp..
CC      Microbial host cells (pref. E. coli) transformed with a recombinant
CC      vector contg. the cDNA, can be used for the prodn. of astaxanthin,
CC      4-ketozeaxanthin, canthaxanthin, echinenone and other
CC      ketocarotenoids. The cDNA was isolated from a cDNA expression
CC      library obtd. from Haematococcus pluvialis NIES-144, in which
CC      astaxanthin biosynthesis was induced by 45 mM acetic acid in
CC      the culture medium.
SQ      Sequence 1629 BP; 348 A; 458 C; 461 G; 362 T;

Query Match 75.0%; Score 15; DB 21; Length 1629;
Best Local Similarity 94.1%; Pred. No. 9.39e+00;

```

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1449 agggaaatcccttcgag 1465
 Cp 18 agggaaatcccttcgag 2

RESULT 15

ID Q26728 standard; DNA; 1860 BP.
 AC Q26728;
 DT 28-JAN-1993 (first entry)
 DE Heat resistant carbamylase gene.
 KW Temp; stable; N-carbamyl-L-amino acid; ss.
 OS Escherichia coli.
 FH Key Location/Qualifiers
 FT cds 320..1550
 FT /*tag= a
 FN J04183391-A.
 PN 30-JUN-1992.
 PD 15-NOV-1990; 307221.
 PP 15-NOV-1990; JP-307221.
 DR (NIPS) NIPPON SODA CO.
 DR WPI; 92-265588/32.
 DR P-PSDB; R25693.
 PT Heat resistant carbamylase - is used to produce L-aminoacid in
 PT high purity by reacting enzyme protein with
 PT N-carbamyl-L-aminoacid
 PS Disclosure; Fig 2; 8pp; Japanese.
 CC The DNA encodes a heat resistant carbamylase (optimum temp. ca. 60
 CC degrees C, optimum pH ca. 8.5) which can be stably produced by
 CC bacteria. The enzyme allows efficient prodn. of a high concn. of N-
 CC carbamyl-L-amino acids which may be hydrolysed to the relavent L-
 CC amino acids in the presence of the heat resistant carbamylase.
 SQ Sequence 1860 BP; 438 A; 456 C; 580 G; 386 T;

Query Match 75.0%; Score 15; DB 4; Length 1860;
 Best Local Similarity 100.0%; Pred. No. 9.39e+00;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 779 gaaggatttccttc 793
 Qy 5 gaaggatttccttc 19

Search completed: Wed May 27 02:58:11 1998
 Job time : 22 secs.

DT 31-MAY-1994 (first entry)
 DE Oligonucleotide probe MK14-A
 KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
 KW ss.
 OS Synthetic.
 PN EP-571911-A.
 PD 01-DEC-1993.
 PF 24-MAY-1993; 108325.
 PR (BECT) BECTON DICKINSON CO.
 PA Shank DD, Spears PA;
 PI WPI; 93-378844/48.
 DR New oligo-nucleotide probes specific for Mycobacteria - used for
 PT detection and amplification of Mycobacteria nucleic acid in
 PT samples
 PS Claim 3; Page 14; 23pp; English.
 CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
 CC (051735). It hybridized to all spp. of mycobacteria tested, but
 CC cross reacted to a few non-mycobacterial spp. The probe may
 CC be useful as an initial screen for mycobacterial infection.
 CC See also Q51735-45 and Q51747-59.

Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 80.0%; Score 16; DB 9; Length 91;
 Best Local Similarity 0.0%; Pred. No. 1.62e+00;
 Matches 0; Conservative 17; Mismatches 1; Indels 0; Gaps 0;

-Db 25 hbsvhhvhhvhsvvv 42

Qy 3 aacttcctaaaggagg 20

RESULT 3
 ID T68207 standard; DNA; 588 BP.
 AC T68207;
 DE Helicobacter pylori-derived open reading frame.
 DE Helicobacter pylori-derived open reading frame.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope; ds.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 cds 1..588
 FT /*tag= a
 FT /note= "no stop codon given"
 PN W09640893-A1.
 PD 19-DEC-1996.
 PF 27-JUN-1995; U09122.
 PR 01-APR-1996; US-487032.
 PA (ASTR) ASTRA AB.
 PI Berglindh OF, Smith D, Mellgaard BL;
 DR WPI; 97-052306/05.
 DR P-PSDB; W20954.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 PT Disclosure; Page 963; 1481pp; English.
 CC The present sequence encodes a Helicobacter pylori-derived protein
 CC of unknown function (no further details given in the specification).
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds.
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.

SQ Sequence 588 BP; 130 A; 123 C; 139 G; 196 T;
 Query Match 75.0%; Score 15; DB 29; Length 588;
 Best Local Similarity 94.1%; Pred. No. 6.18e+00;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Db 494 tcccttttaggaattc 510
 Cp 18 tcccttttaggaattc 2
 RESULT 4
 ID Q40312 standard; DNA; 3560 BP.
 AC Q40312;
 DT 03-AUG-1993 (first entry)
 DE Helicobacter pylori urease operon.
 KW ureE; ureF; ureG; ureH; ureI; stomach; gastric mucosa; gastric ulcer;
 KW duodenal ulcer; acid tolerance; ss.
 OS Helicobacter pylori.
 FH Key Location/Qualifiers
 rbs 198..201
 FT /*tag= a
 FT /standard_name= Shine_Dalgarno
 FT cds 212..799
 FT /*tag= b
 FT /label= ureI
 FT cds 801..1313
 FT /*tag= c
 FT /label= ureE
 FT rbs 1313..1316
 FT /*tag= d
 FT /standard_name= Shine_Dalgarno
 FT cds 1325..2095
 FT /*tag= e
 FT /label= ureF
 FT misc_feature 1819..1834
 FT /*tag= f
 FT /note= "sigma-54 promoter-like sequence"
 FT cds 2123..2722
 FT /*tag= g
 FT /label= ureG
 FT rbs 2713..2717
 FT /*tag= h
 FT /standard_name= Shine_Dalgarno
 FT cds 2722..3519
 FT /*tag= i
 FT /label= ureH
 FT terminator 66..89
 FT /*tag= j
 FT /note= "rho-independent hairpin"
 FT rbs 2111..2116
 FT /*tag= k
 FT /standard_name= Shine_Dalgarno
 FT terminator 3528..3560
 FT /*tag= l
 FT /note= "rho-independent hairpin"
 PN W09307273-A.
 PD 15-APR-1993.
 PF 02-OCT-1992; F00921.
 PR 03-OCT-1991; FR-012198.
 PA (INRM) INSERM INST NAT SANTE & RECH MED.
 PI (INSP) INST PASTEUR.
 PI Cussac V, Ferero R, Labigne A;
 DR WPI; 93-134462/16.
 DR P-PSDB; R34395, R34396, R34397, R34398, R34399.
 PT Helicobacter pylori genes useful in diagnosis, vaccines and
 PT treatment - necessary for the regulation and maturation of urease
 PS Claim 7-11; Fig 4; 94pp; French.
 CC The five new urease genes ure E, F, G, H and I were identified by
 CC deletion studies in E. coli where all 5 were found to be necessary
 CC for functional urease expression. They encode proteins of
 CC estimated mol. wt. 19.5, 28.6, 21.7, 29.6 and 21.7kD, respectively.
 CC Recombinant H. pylori having a mutation in at least one of the 5 new

W P E R E H

(TM)

Release 3.0.5AA John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1997 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

MP: Wed May 27 02:59:48 1998; Mapar time 16.53 Seconds
154.103 Million cell updates/sec
Tabular output not generated.

Title: >SEQ2
Description: (1-20) from new.seq
Perfect Score: 20
N.A. Sequence: 1 ggaacttccttaaggagg 20
Comp: ccttgaaggattccctcc

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0
Searched: 176923 seqs, 63680241 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-genseq1-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37

Statistics: Mean 5.133; Variance 2.741; scale 1.873

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
C 1	18	90.0	91	Q51746	Oligonucleotide probe	1.01e-01
C 2	16	80.0	91	Q51746	Oligonucleotide probe	1.62e+00
C 3	15	75.0	588	29	Helicobacter pylori-d	6.18e+00
C 4	15	75.0	3560	6	Helicobacter pylori u	6.18e+00
C 5	14	70.0	780	1	Probe PCS.7 for scree	2.28e+01
C 6	14	70.0	1041	29	H. pylori inner membr	2.28e+01
C 7	14	70.0	1065	29	H. pylori cell envelo	2.28e+01
C 8	14	70.0	1074	29	H. pylori cell envelo	2.28e+01
C 9	14	70.0	1404	7	Sequence which hybrid	2.28e+01
C 10	14	70.0	3134	17	b98 insert encoding a	2.28e+01
C 11	14	70.0	4488	8	Human FACC cDNA clone	2.28e+01
C 12	14	70.0	5342	22	DNA polymerase.	2.28e+01
C 13	14	70.0	5342	22	DNA polymerase gene,	2.28e+01
C 14	14	70.0	19932	29	CagI locus.	2.28e+01

C 15	13	65.0	55	37	V03751	Porcine TNFalpha-conv	8.10e+01
C 16	13	65.0	89	37	V03750	Porcine TNFalpha-conv	8.10e+01
C 17	13	65.0	266	8	Q59163	Human brain Expressed	8.10e+01
C 18	13	65.0	266	6	Q39751	Expressed Sequence Ta	8.10e+01
C 19	13	65.0	273	18	T19188	Human gene signature	8.10e+01
C 20	13	65.0	306	8	O60351	Human brain Expressed	8.10e+01
C 21	13	65.0	372	8	O60269	Human brain Expressed	8.10e+01
C 22	13	65.0	1003	30	T40041	Presenilin-1 exon 12.	8.10e+01
C 23	13	65.0	1608	22	T40731	Endonuclease 2 encodi	8.10e+01
C 24	13	65.0	1655	7	Q43720	Sequence which hybrid	8.10e+01
C 25	13	65.0	1751	1	N71021	Sequence encoding D-a	8.10e+01
C 26	13	65.0	2157	9	O56926	Human TGF-beta-3.	8.10e+01
C 27	13	65.0	2158	1	O02820	cDNA sequence encodin	8.10e+01
C 28	13	65.0	2464	37	V03752	Porcine TNFalpha-conv	8.10e+01
C 29	13	65.0	2529	3	Q20576	Transforming Growth F	8.10e+01
C 30	13	65.0	2530	2	Q06845	Sequence encoding tum	8.10e+01
C 31	13	65.0	2861	23	T28639	RNase L inhibitor "H2	8.10e+01
C 32	13	65.0	3568	23	T28638	RNase L inhibitor "RL	8.10e+01
C 33	13	65.0	4080	20	T30652	P. gingivalis haemagg	8.10e+01
C 34	13	65.0	4998	1	N81114	Non-A, non-B hepatitis	8.10e+01
C 35	13	65.0	5340	31	T71296	KOD1 thermostable DNA	8.10e+01
C 36	13	65.0	5342	22	T28360	DNA polymerase gene,	8.10e+01
C 37	13	65.0	5342	22	T14694	DNA polymerase.	8.10e+01
C 38	13	65.0	6000	33	T78851	Lys-specific thiol en	8.10e+01
C 39	13	65.0	6241	20	T30655	P. gingivalis haemagg	8.10e+01
C 40	13	65.0	6895	20	T30653	P. gingivalis porphy	8.10e+01
C 41	13	65.0	7266	14	Q83489	Arg-gingipain-2 gene.	8.10e+01
C 42	13	65.0	8640	20	T30656	P. gingivalis haemagg	8.10e+01
C 43	13	65.0	12537	24	T41705	Lymphocyte specific i	8.10e+01
C 44	13	65.0	18177	1	N90490	DNA of human retinobl	8.10e+01
C 45	13	65.0	580073	27	T58840	Mycoplasma genitalium	8.10e+01

ALIGNMENTS

RESULT 1
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PR 24-MAY-1993; 108325.
PF 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI; 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 90.0%; Score 18; DB 9; Length 91;
Best Local Similarity 0.0%; Pred. No. 1.01e-01;
Matches 0; Conservative 19; Mismatches 1; Indels 0; Gaps 0;

Db 27 svhhvhhvhsvvvhhv 46

Cp 20 cctcccttaggaagttcc 1

RESULT 2
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;

TITLE Wan,K.H., Whitelaw,K.R., Yee,A., Yeh,R.T., Yu,C. and Palazzolo,M.J.
JOURNAL Direct Submission
COMMENT Submitted (09-APR-1997)
Sequence submitted by:
Human Genome Center
Lawrence Berkeley National Laboratory, MS 74-157
Berkeley, CA 94720

For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Website (<http://www-hgc.lbl.gov/sequence-archive.html>) or
send email to human@genome.lbl.gov.

FEATURES

source

1. .3304
/location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="Pl H69 (5403)"
/sub_clone="1a3"

BASE COUNT 1091 a 639 c 602 g 972 t

ORIGIN

Very Match 75.08; Score 15; DB 22; Length 3304;

Best Local Similarity 94.18; Pred.No. 5.48e+01;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 2082 ACTTCCCAAGGAGG 2098

||||| |||||||

.Qy 4 acttcctaaaggagg 20

RESULT 15

LOCUS

HECURES 3560 bp DNA BCT 11-MAY-1994

DEFINITION

Helicobacter pylori ureE, ureF, ureG, ureH, and ureI gene, complete

ACCESSION

M84338

NID g149012

KEYWORDS

UREE gene; UREH gene; UREG gene; ureH gene; ureI gene.

SOURCE

Helicobacter pylori (individual isolate 85P) DNA.

ORGANISM

Helicobacter pylori

REFERENCE

1 (bases 1 to 3560)

AUTHORS

Cussac,V., Ferrero,R. and Labigne,A.F.

TITLE

Expression of Helicobacter pylori urease genes in Escherichia coli

JOURNAL

grown under nitrogen-limiting conditions

MEDLINE

J. Bacteriol. 174, 2466-2473 (1992)

FEATURES

source

1. .3560
/location/Qualifiers

/organism="Helicobacter pylori"

/isolate="85P"

/db_xref="taxon:210"

RBS

197. .204

gene

/standard_name="Shine-Dalgarno site"

CDS

/note="putative"

212. .799

/gene="ureI"

212. .799

/gene="ureI"

/codon_start=1

/db_xref="PID:g485332"

/transl_table=11

/translation="MLGLVLYVGVILNSIGLTKVDPKSTAVMNFVGLSIICN

VVITYSALNPAPYVGADIAQVSHLNFYCPATGLLFGTYLYAAINHFGLDWR

PYSWISFVAINTIPAAILSHSDMLDGHKVLGITEGDWWAILWANGVILWLTAFIEN

ILKIPLGKTPMLAIEGILTAWIPAWLLFIQHW"

801. .1313

/gene="ureE"

801. .1313

/gene="ureE"

/codon_start=1

/db_xref="PID:g485333"

/transl_table=11

/translation="MIERLIGNRLNPLDFSDYVDLEWFETFRKKIARFKTRQCKD
IAVRUKDAPKLGFSOGDILFKKEKEIIAVNILDSEVIHOAKSVAEVAKICVEIGNRH
AALYGESQEFKTFEPPTLALLEKLGQVNRVLSKLDSEKRLTVSMPSPEPKVS
IASDFKVMK"
1312. .1316
/standard_name="Shine-Dalgarno site"

RBS

/note="putative"

gene

1325. .2095

CDS

/gene="ureF"

1325. .2095

/gene="ureF"

/codon_start=1

/db_xref="PID:g485334"

/transl_table=11

/translation="MDKGSVKSIKSVGMPLPKPTDSNAHVDNEFLLOVNDVFP

IGSYTHSPGLLARNLHPAKVTKNSALKYLKANLSQFLYTEMLSKLTYESALQQD

LKRILGVEIITLSTSPHMLRANOKLGNRFKTKLOANNELDIGAFFNAYAQOTEDPT

HATSYGVAASGLIELKALRHLRYAOTSNMVINCQKSVPLUSQNDGQKILLLSQSPFN

OLIENTLEDESHLCAASVQNDIKAMQHESLYSRLYS"

2112. .2117

/standard_name="Shine-Dalgarno site"

/note="putative"

2124. .2723

/gene="ureG"

2124. .2723

/gene="ureG"

/codon_start=1

/db_xref="PID:g485335"

/transl_table=11

/translation="MVKIGVCGVSGKGTALIEALTRHMSKDYDMVITNDIYTKEDA

EFMCKNSVMPREIIGVETGCPHTAIREDASMNLEAVEEMHGRFPNLELLIESGGS

NLSATFNPELADFTIFVIDVAEGDKIPKGGPGITRSDLLVINKIDLAPYVGADLKVM

ERDSKKIAAKSPFLPNIRAKEGLDDVIAWIKRNALLD"

2714. .2719

/gene="ureG"

/standard_name="Shine-Dalgarno site"

/note="putative"

2723. .3520

/gene="ureH"

2723. .3520

/gene="ureH"

/codon_start=1

/db_xref="PID:g485336"

/transl_table=11

/translation="MNTVAQESKRLKTKIGADRCVIEDNDFPFPFKIMAFYPKDD

LAEMLLAVSPGLMKGAQDVQLNIGPNCKLRITQSQFEKIHNTEDGFASRDMHIVVG

ENAFIDFAPPLIPFENAHFNGNTTISLSSQLLYSEIIVAGRVARNELEKFNRLHT

KISILQDERPIYDNTILDPKTDNNMCMFDGYTHYLNVLVNCPIELSGVRGLIEE

SEGVDGAVSEIASSHLCLAKAGSEPLLHLUREKIARFITQITPKV"

BASE COUNT 1060 a 698 c 777 g 1025 t

ORIGIN

Query Match 75.08; Score 15; DB 13; Length 3560;

Best Local Similarity 94.18; Pred.No. 5.48e+01;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 702 TCCTTTTAGGGAATTC 718

||||| ||||||| |||

Cp 18 tccttttagggaagttc 2

Search completed: Wed May 27 02:59:30 1998

Job time : 61 secs.

Query Match 75.0%; Score 15; DB 27; Length 557;
Best Local Similarity 88.9%; Pred. No. 5.48e+01;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 335 AACTTCCTCAAGGNGG 352
||||| ||||| ||||| ||
Oy 3 aacttcctaaaggagg 20

RESULT 12
LOCUS GGAX1EX1 1398 bp DNA VRT 07-FEB-1997

DEFINITION G.gallus axonin-1 gene, exon 1.
ACCESSION X79608
NID 9535164
KEYWORDS axonin-1; exon; promoter.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
vertebrates; Archosauria; Aves; Neognathae; Galliformes;
Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1398)
Giger,R.J., Vogt,L., Zuellig,R.A., Rader,C., Hennehan-Beatty,A.,
Wolfer,D.P. and Sonderegger,P.
The gene of chicken axonin-1. Complete structure and analysis of
the promoter
Eur. J. Biochem. 227 (3), 617-628 (1995)
MEDLINE 95172044
REFERENCE 2 (bases 1 to 1398)
Giger,R.J.
Direct Submission
Submitted (08-JUN-1994) R.J. Giger, University of Zuerich,
Biochemisches Institut, Winterthurerstr 190, CH-8057 Zuerich,
SWITZERLAND

FEATURES
source
1. 1398
/organism="Gallus gallus"
/db_xref="taxon:9031"
/dev_stage="adult"
/tissue_type="liver"
/clone_lib="lambda EMBL-3"
1127..1232
/gene="axonin-1"
/number=1
1127..1398
/gene="axonin-1"
1233..>1398
/gene="axonin-1"
/number=1
BASE COUNT 326 a 426 c 406 g 240 t
ORIGIN

Query Match 75.0%; Score 15; DB 17; Length 1398;
Best Local Similarity 94.1%; Pred. No. 5.48e+01;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1335 AACTTCCTCATAGGAG 1351
||||| ||||| ||||| ||
Oy 3 aacttcctaaaggagg 19

RESULT 13
LOCUS ASU78028 1416 bp DNA BCT 07-MAY-1997

DEFINITION Arthrobacter sp. beta-galactosidase gene, complete cds.
ACCESSION U78028
NID 91857332
KEYWORDS Arthrobacter sp.
SOURCE Arthrobacter sp.
ORGANISM Eubacteria; Firmicutes; Actinomycetes; Arthrobacter.
REFERENCE 1 (bases 1 to 1416)

AUTHORS Gutshall,K., Wang,K. and Brenchley,J.E.
TITLE A novel Arthrobacter beta-galactosidase with homology to eucaryotic
beta-galactosidases
J. Bacteriol. 179 (9), 3064-3067 (1997)
MEDLINE 97284517
REFERENCE 2 (bases 1 to 1416)
Gutshall,K., Wang,K. and Brenchley,J.
Direct Submission
TITLE Submitted (12-NOV-1996) Biochemistry and Molecular Biology, Penn
State University, 211 South Frear Lab, University Park, PA 16802,
USA

FEATURES
source
1. 1416
/organism="Arthrobacter sp."
/db_xref="taxon:1667"
/clone="B7-14"
1. 1416
/note="family 35 beta-galactosidase"
/codon_start=1
/product="beta-galactosidase"
/db_xref="PID:g1857333"
/transl_table=1
/translation="MSVFTIGEQDFLNDQPHRLAGALHYFRVHPDQWADRIRKARQ
MGLTETIYVANNLHAPSEDVFTSAGLDLGRFLDLVAEGMHAIVRPGPYICAEWDN
GGLPGLWFKSGNPVIRTSDPVYALVSRYNEALAPILVPRQIDRGGPIILVOIENEYK
AYGSDHYLEQLVELNREIGLSVFFTRGSRQPEPDADQKQARTCTROPDSVESOR
NARPCASHPTGATHVIGEGFAGFEGPLGDHHTTSVQESVHELELLAAGASVNV
YMFHGGTFNGMNGANDKGYQPTVTSYDADFLDEAGQPTKYIWAFFREVLGRYELV
DQIDIPAPRQAGQEFVSLQDLMFWEYLDATASWTQSDPLTTDAVGSFNGFSYRSQ
ISVGEOLRAQLFRGAGDLPEPRAGHPLPRSEGRDGTARRPPPEPWKLCSSRIKA
GSTTGKGHWARLRDSSAPPHQ"
BASE COUNT 301 a 422 c 418 g 275 t
ORIGIN

Query Match 75.0%; Score 15; DB 13; Length 1416;
Best Local Similarity 100.0%; Pred. No. 5.48e+01;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 965 CCTTTAGGGAAGTTC 979
||||| ||||| ||||| ||
Cp 16 cctttagggaagttc 2

RESULT 14
LOCUS HSL81905 3304 bp DNA PRI 09-APR-1997

DEFINITION Homo sapiens (subclone 1_a3 from P1 H69) DNA sequence, complete
sequence.
ACCESSION L81905
NID g1930225
KEYWORDS HTG.
SOURCE Homo sapiens (Subclones in pOT2 from P1 clone H69) DNA.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 3304)
Martin,C.H., Arcaina,T., Bondoc,M.M., Chiang,A., Critz,P.A.,
Davis,C.A., Doyle,C.M., Ericsson,C.L., Farfan,D.E., Gunning,K.M.,
Houston,K.A., Jaklevic,M.A., Kadner,K.E., Kim,K., Kim,S.F.,
Ko,C.L., Lewis,K.D., Li,M., Lindquist,K.J., Lomotan,M.A.,
Lustre,V.M., Machrus,M.U., Mayeda,C.A., Miguel,T.M., Miller,C.A.,
Mok,M.S., Pacleb,J.M., Patel,S.G., Santos,R.F., Subramanian,S.,
Wan,K.H., Whitelaw,K.R., Yee,A., Yeh,R.T., Yu,C. and Palazzolo,M.J.
Sequencing of human chromosome 5q
Unpublished (1996)
2 (bases 1 to 3304)
Martin,C.H., Arcaina,T., Bondoc,M.M., Chiang,A., Critz,P.A.,
Davis,C.A., Doyle,C.M., Ericsson,C.L., Farfan,D.E., Gunning,K.M.,
Houston,K.A., Jaklevic,M.A., Kadner,K.E., Kim,K., Kim,S.F.,
Ko,C.L., Lewis,K.D., Li,M., Lindquist,K.J., Lomotan,M.A.,
Lustre,V.M., Machrus,M.U., Mayeda,C.A., Miguel,T.M., Miller,C.A.,
Mok,M.S., Pacleb,J.M., Patel,S.G., Santos,R.F., Subramanian,S.,
Wan,K.H., Whitelaw,K.R., Yee,A., Yeh,R.T., Yu,C. and Palazzolo,M.J.

Best Local Similarity 94.4%; Pred. No. 1.29e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 226937 GGAACCTCCCTAAAGTGA 226954
|||||
QY 1 ggaactccctaaaggga 18

RESULT 9 BB58SRR 533 bp RNA PLN 08-JAN-1997
LOCUS

DEFINITION B.brongniartii 5.8S rRNA gene and internal transcribed spacers 1 and 2.
ACCESSION 254103
NID g1769928
KEYWORDS 5.8S ribosomal RNA; 5.8S rRNA gene; internal transcribed spacer; ITS2.
internal transcribed spacer 1; internal transcribed spacer 2; ITS1; ITS2.

SOURCE
ORGANISM Beauveria brongniartii.
Eukaryotae; Mitochondrial eukaryotes; Fungi; Ascomycota; mitosporic
Ascomycota; Beauveria.
REFERENCE 1 (bases 1 to 533)
AUTHORS Shih, H., Yuan, G., and Tzean, S.
TITLE The phylogeny of Beauveria spp. based on 5.8S rDNA and flanking
internal transcribed spacers
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 533)
AUTHORS Tzean, S.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1995) Tzean S., National Taiwan University,
Plant pathology and Entomology, 1, Roosevelt Rd. Sec. 4, Taipei,
Taiwan, Republic of China, 107

FEATURES
source Location/Qualifiers
1..533
/organism="Beauveria brongniartii"
/isolate="CCRC32838(-CBS223.53)"
/db_xref="taxon:37993"
/tissue_type="mycelium"
11..168
/note="internal transcribed spacer I (ITS1)"
/citation=[1]
169..325
/gene="5.8S rRNA"
/citation=[1]
/product="5.8S ribosomal RNA"
169..325
/gene="5.8S rRNA"
326..527
/note="internal transcribed spacer II (ITS2)"
/citation=[1]
BASE COUNT 133 a 158 c 133 g 109 t
ORIGIN

Query Match 75.0%; Score 15; DB 20; Length 533;
Best Local Similarity 94.1%; Pred. No. 5.48e+01;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 341 CCTCCCTTTGGGGAAGT 357
|||||
Cp 20 cctccctttagggaagt 4

RESULT 10 BT58SRR 537 bp RNA PLN 08-JAN-1997
LOCUS

DEFINITION B.tenella 5.8S rRNA gene and internal transcribed spacers 1 and 2.
ACCESSION 254107
NID g1770083
KEYWORDS 5.8S ribosomal RNA; 5.8S rRNA gene; internal transcribed spacer; ITS2.
internal transcribed spacer 1; internal transcribed spacer 2; ITS1; ITS2.
Beauveria tenella.

ORGANISM Beauveria tenella
Eukaryotae; Mitochondrial eukaryotes; Fungi; Ascomycota; mitosporic
Ascomycota; Beauveria.
REFERENCE 1 (bases 1 to 537)
AUTHORS Shih, H., Yuan, G., and Tzean, S.
TITLE The phylogeny of Beauveria spp. based on 5.8S rDNA and flanking
internal spacers
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 537)
AUTHORS Tzean, S.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1995) Tzean S., National Taiwan University,
Plant pathology and Entomology, 1, Roosevelt Rd. Sec. 4, Taipei,
Taiwan, Republic of China, 107

FEATURES
source Location/Qualifiers
1..537
/organism="Beauveria tenella"
/db_xref="taxon:37999"
/tissue_type="mycelium"
11..171
/note="internal transcribed spacer I (ITS1)"
/citation=[1]
172..326
/gene="5.8S rRNA"
/citation=[1]
/product="5.8S ribosomal RNA"
172..326
/gene="5.8S rRNA"
327..531
/note="internal transcribed spacer II (ITS2)"
/citation=[1]
BASE COUNT 135 a 159 c 130 g 113 t
ORIGIN

Query Match 75.0%; Score 15; DB 20; Length 537;
Best Local Similarity 94.1%; Pred. No. 5.48e+01;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 341 CCTCCCTTTGGGGAAGT 357
|||||
Cp 20 cctccctttagggaagt 4

RESULT 11 SVU27968 557 bp DNA VRL 23-JUL-1995
LOCUS

DEFINITION Stealth virus clone C16257 T7.
ACCESSION U27968
NID 9903573
KEYWORDS Stealth virus.
SOURCE Stealth virus.
ORGANISM Viruses; GSDNA viruses, no RNA stage; Herpesviridae;
Betaherpesvirinae; unclassified Betaherpesvirinae.
REFERENCE 1 (bases 1 to 557)
AUTHORS Martin, W.J., Zeng, L.C., Ahmed, K., and Roy, M.
TITLE Cytomegalovirus-related sequence in an atypical cytopathic virus
repeatedly isolated from a patient with chronic fatigue syndrome
Am. J. Pathol. 145 (2), 440-451 (1994)

JOURNAL 94330517
MEDLINE 2 (bases 1 to 557)
REFERENCE Martin, W.J.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-1995) W. John Martin, Pathology, University
Diagnostic Laboratory, 1840 North Soto Street, Los Angeles, CA
90033, USA

FEATURES
source Location/Qualifiers
1..557
/organism="Stealth virus"
/db_xref="taxon:36452"
/clone="C16257 T7"
BASE COUNT 124 a 135 c 126 g 114 t 58 others
ORIGIN

Howard,S., Jerome,N., Koop,B.F., Lee,H., Loretz,C., Paepel,B.,
 Zackrone,K. and Hood,L.
 Sequence determination of the human T cell receptor beta locus:
 Strategy and error analysis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 267156)
 AUTHORS Rowen,L., Wang,K., Boysen,C., Ahearn,M.E., Charmley,P., Paepel,B.,
 Lee,I., Chen,L., Trask,B., Nickerson,D., Seto,D. and Hood,L.
 TITLE Sequence variation among several haplotypes in the human T cell
 receptor beta locus
 JOURNAL Unpublished
 REFERENCE 4 (bases 1 to 267156)
 AUTHORS Rowen,L.
 TITLE Direct Submission
 JOURNAL Submitted (15-OCT-1994) L.Rowen leerowen@u.washington.edu
 REFERENCE 5 (bases 1 to 267156)
 AUTHORS Rowen,L.
 TITLE Direct Submission
 JOURNAL Submitted (24-JUN-1997)
 COMMENT This sequence overlaps section 2 of the T cell receptor beta locus,
 Genbank Accession Number U66060, by 9647 bases.
 FEATURES
 source
 1. .267156
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="7q35"
 source
 1. .37031
 /note="(vector PWE15A)"
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="CGM1, haplotype B"
 /clone_lib="YAC D49H4"
 /germline
 complement(1. .240)
 /note="putative"
 /rpt_family="Alu"
 repeat_region
 complement(577. .709)
 /note="SINE; putative"
 /rpt_family="MIR"
 repeat_region
 complement(1463. .1575)
 /note="MaLR retroposon LTR; putative"
 /rpt_family="MLTIG"
 repeat_region
 2307. .2428
 /note="DNA transposon fossil; putative"
 /rpt_family="MER33; fragment 1"
 2429. .2988
 /note="putative"
 /rpt_family="LINE 1"
 repeat_region
 2429. .7244
 /note="LINE; putative"
 /rpt_family="L1PA7; fragment 1"
 7245. .7975
 /note="LINE; putative"
 /rpt_family="L1PA7; fragment 2"
 repeat_region
 7997. .8234
 /note="putative"
 /rpt_family="Alu"
 8295. .8463
 /note="putative"
 /rpt_family="MER33"
 repeat_region
 8784. .8793
 /note="15 bp tandem duplication; putative"
 /rpt_type=tandem
 repeat_region
 complement(9085. .9224)
 /note="DNA transposon fossil; putative"
 /rpt_family="MER63"
 9888. .9965
 /note="SINE; putative"
 /rpt_family="MIR"
 10076. .10320
 /note="LINE; putative"
 /rpt_family="L1MA4; fragment 1"

repeat_region
 10321. .10616
 /note="putative"
 /rpt_family="Alu"
 repeat_region
 10617. .10747
 /note="LINE; putative"
 /rpt_family="L1MA4; fragment 2"
 12541. .13965
 /note="60 (120) mer, 19x repeated, 91x conserved first
 5.5 120 mers; putative"
 /rpt_type=tandem
 repeat_region
 15873. .16381
 /note="DNA transposon fossil; putative"
 /rpt_family="MER64"
 complement(19303. .19385)
 /note="SINE; putative"
 /rpt_family="MIR"
 19542. .19632
 /note="SINE; putative"
 /rpt_family="MIR"
 repeat_region
 19682. .19755
 /note="LINE; putative"
 /rpt_family="L1MA9"
 19945. .20437
 /note="retroposon LTR; putative"
 /rpt_family="MER5B"
 complement(21066. .21541)
 /note="retroposon LTR; putative"
 /rpt_family="MER60A"
 complement(22822. .23202)
 /note="LINE; putative"
 /rpt_family="L1PA16"
 repeat_region
 23444. .23601
 /note="SINE; putative"
 /rpt_family="MIR"
 complement(23615. .23909)
 /note="putative"
 /rpt_family="Alu"
 repeat_region
 24043. .24079
 /note="SINE; putative"
 /rpt_family="MIR"
 complement(24304. .24697)
 /note="endogenous retroviral LTR; putative"
 /rpt_family="LTR1; fragment 1"
 complement(25086. .25502)
 /note="endogenous retroviral LTR; putative"
 /rpt_family="LTR1; fragment 2"
 complement(25975. .26233)
 /note="putative"
 /rpt_family="Alu"
 repeat_region
 27165. .28630
 /note="LINE; putative"
 /rpt_family="L1P5; fragment 1"
 28056. .28507
 /note="putative"
 /rpt_family="LINE 1"
 repeat_region
 28631. .28702
 /note="LINE; putative"
 /rpt_family="L1P5; fragment 2"
 complement(28713. .29101)
 /note="LINE; putative"
 /rpt_family="mir2"
 repeat_region
 complement(30772. .31746)
 /note="LINE; putative"
 /rpt_family="MER42C"
 repeat_region
 complement(31996. .32116)
 /note="DNA transposon fossil; putative"
 /rpt_family="MER5B"
 repeat_region
 32313. .32419
 /note="LINE; putative"

...
 Note: remainder of annotations omitted.

Query Match 80.08; Score 16; DB 22; Length 267156;

*by a single finished sequence with the same accession number.

```

* 1 2753: contig of 2753 bp in length
* 2754 3442: gap of unknown length
* 3443 6405: contig of 2963 bp in length
* 6406 1793: gap of unknown length
* 1794 11354: contig of 4261 bp in length
* 11355 12042: gap of unknown length
* 12043 17983: contig of 5941 bp in length
* 17984 18671: gap of unknown length
* 18672 49612: contig of 30941 bp in length
* 49613 50300: gap of unknown length
* 50301 102000: contig of 51700 bp in length.

```

FEATURES

source

```

1..102000
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"

```

BASE COUNT 31365 a 17174 c 17956 g 32064 t 3441 others

ORIGIN

```

Query Match 80.0%; Score 16; DB 14; Length 102000;
Best Local Similarity 100.0%; Pred. No. 1.29e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 49403 TCCCTTTAGGAGGTT 49418

```

|||||
Cp 18 tccctttagggaagtt 3

```

RESULT

LOCUS

7 AB000882 148960 bp DNA

PRI

04-APR-1998

DEFINITION Homo sapiens genomic DNA for centromeric end of MHC class I region on chromosome 6, complete sequence.

ACCESSION

AB000882

NID

93021697

KEYWORDS

HTG.

SOURCE

Homo sapiens (haplotype:A2 B62 CW10 DR4) adult male blood

immunoresponce cell cell_line:BOLETH DNA, clone_lib:CEPH YAC

clone:745D12.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominoidea; Homo.

1 (sites)

Shina.T., Tamiya.G., Oka.A., Yamagata.T., Yamagata.N., Kikkawa.E.,

Goto.K., Mizuki.N., Watanabe.K., Fukuzumi.Y., Taguchi.S.,

Sugawara.C., Ono.A., Chen.L., Yamazaki.M., Fashiro.H., Ando.A.,

Ikemura.T., Kimura.M. and Inoko.H.

Nucleotide sequencing analysis of the 146-kilobase segment around

the IkbL and MICA genes at the centromeric end of the HLA class I

region

Genomics 47 (3), 372-382 (1998)

98149985

2 (bases 1 to 148960)

Shina.T.

Direct Submission

Submitted (04-FEB-1997) to the DDBJ/EMBL/GenBank databases. Takashi

Shina, Tokai University School of Medicine, Department of

Molecular Life Sciences; Bonseidai, Isehara, Kanagawa 259-11, Japan

(E-mail: tshina@is.icc.u-tokai.ac.jp, Tel: 0463-93-1121,

Fax: 0463-94-8884)

Location/Qualifiers

1..148960

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="BOLETH"

/cell_type="immunoresponce cell"

/chromosome="6"

/clone="745D12"

/clone_lib="CEPH YAC"

/dev_stage="adult"

/haplotype="A2 B62 CW10 DR4"

/map="6p21.3"

/sex="male"

/tissue_type="blood"

BASE COUNT 41138 a 33995 c 34022 g 39805 t

ORIGIN

Query Match 80.0%; Score 16; DB 22; Length 148960;

Best Local Similarity 94.4%; Pred. No. 1.29e+01;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 87931 TCCCTTAAGGAGGTTCC 87948

```

|||||
Cp 18 tccctttagggaagttcc 1

```

FEATURES

source

8 U66059 267156 bp DNA

PRI

22-JUL-1997

DEFINITION

Human germline T-cell receptor beta chain

Dopamine-beta-hydroxylase-like, TRY1, TRY2, TRY3, TCRBV27SLP,
TCRBV22SIA2NIT, TCRBV9S1A1T, TCRBV9S2A2PT, TCRBV5S1A1T, TCRBV13S3,
TCRBV6S7P, TCRBV7S3A2T, TCRBV13S2A1T, TCRBV9S2A2PT, TCRBV7S2A1N4T,
TCRBV13S9/13S2A1T, TCRBV6S5A1N1, TCRBV30S1P, TCRBV31S1, TCRBV13S5,
TCRBV6S1A1N1, TCRBV32S1P, TCRBV5S5P, TCRBV1S1A1N1, TCRBV12S2A1T,
TCRBV21S1, TCRBV8S4P, TCRBV12S3, TCRBV21S3A2N2T, TCRBV8S5P,
TCRBV13S1 genes from bases 1 to 267156 (section 1 of 3).

ACCESSION

U66059

NID

L36092

KEYWORDS

G1552494

SOURCE

C region; C-beta gene segment; D region; J-beta gene segment;

J-segment; T-cell receptor beta-chain; TCR-beta gene; V-beta gene

segment; V-segment; cell membrane protein; constant region;

diversity region; germline; joining segment; tryptsin; trypsinogen;

variable segment.

Homo sapiens (clone: K41A) (clone_lib: CGM1; YAC D49H4) DNA; Homo

sapiens (clone: K35) (clone_lib: CGM1; YAC D49H4) DNA; Homo sapiens

(clone: K26) (clone_lib: CGM1; YAC D49H4) DNA; Homo sapiens (clone: K56)

(clone_lib: CGM1; YAC D49H4) DNA; Homo sapiens (library: ATCC

1521) (clone: X21B) (clone_lib: Kai Wang's) DNA; Homo sapiens

(library: ATCC 1521) (clone: G54) (clone_lib: Kai Wang's) DNA; Homo

sapiens (library: HeLa) (clone: H137) (clone_lib: Eric Lai's) DNA;

Homo sapiens (library: sperm) (clone: H18) (clone_lib: Eric Lai's)

DNA; Homo sapiens (clone: H18/G15gap) DNA; Homo sapiens (library:

ATCC 1521) (clone: G15) (clone_lib: Kai Wang's) DNA; Homo sapiens

(library: ATCC 1521) (clone: X1A) (clone_lib: Kai Wang's) DNA; Homo

sapiens (clone: A27) (clone_lib: CGM1; YAC 234 A6F6) DNA; Homo

sapiens (clone: A212partial) (clone_lib: CGM1; YAC 234 A6F6) DNA;

Homo sapiens (clone: A14) (clone_lib: CGM1; YAC 234 A6F6) DNA;

sapiens (library: sperm) (clone: H7.1) (clone_lib: Eric Lai's) DNA;

Homo sapiens (clone: H12.18) (clone_lib: Eric Lai's) DNA; Homo

sapiens (library: HeLa) (clone: H130) (clone_lib: Eric Lai's) DNA;

sapiens (library: HeLa) (clone: H130) (clone_lib: Eric Lai's) DNA;

Homo sapiens (clone: A16) (clone_lib: CGM1; YAC 234 A6F6) DNA; Homo

sapiens (clone: C215) (clone_lib: CGM1; YAC 234 A72B3) DNA; Homo

sapiens (clone: G1) (library: ATCC 1521) (clone_lib: Kai Wang's)

DNA; Homo sapiens (clone: C68) (clone_lib: CGM1; YAC 210 A38G1)

DNA; Homo sapiens (clone: C21) (clone_lib: CGM1; YAC 234 A72B3)

DNA; Homo sapiens (library: ATCC 1521) (clone: X11) (clone_lib: Kai

Wang's) DNA; Homo sapiens (library: ATCC 1521) (clone: X6A)

(clone_lib: Kai Wang's) DNA; Homo sapiens library: ATCC 1521)

(clone: CBG1) (clone_lib: Kai Wang's) DNA; Homo sapiens (clone:

CBG1/C299gap) DNA; and Homo sapiens (library: ATCC 1521) (clone:

C29) (clone_lib: Kai Wang's) DNA.

ORGANISM

Homo sapiens

Eukaryota; Eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;

Homo.

REFERENCE

1 (bases 1 to 267156)

Rowen.L., Koop,B.F. and Hood,L.

The complete 685-kilobase DNA sequence of the human beta T cell

receptor locus

Science 272 (5269), 1755-1762 (1996)

JOURNAL

MEDLINE

66256474

REFERENCE

2 (bases 1 to 267156)

Rowen.L., Seto,J., Smit,A., Acharya,C., Ahearn,M.E., Ankener,M.,

Baskin,D., Bumgarner,R., Chen,L., Chen,N., Deshpande,P., Faust,J.,

```

DEFINITION Homo sapiens genomic DNA for centromeric end of MHC class I region
on chromosome 6, cosmid clone: TY3A9, complete sequence.
ACCESSION AB000879
NID 93021695
KEYWORDS HTG.
SOURCE Homo sapiens (haplotype:A2 B62 CW10 DR4) adult male blood
immunoresponse cell cell_line:BOLETH DNA, clone_lib:CEPH YAC
clone:745D12 sub_clone:cosmid clone:TY3A9.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Shihina,T., Tamiya,G., Oka,A., Yamagata,T., Yamagata,N., Kikkawa,E.,
Goto,K., Mizuki,N., Watanabe,K., Fukuzumi,Y., Taguchi,S.,
Sugawara,C., Ono,A., Chen,L., Yamazaki,M., Tashiro,H., Ando,A.,
Ikemura,T., Kimura,M. and Inoko,H.
TITLE Nucleotide sequencing analysis of the 146-kilobase segment around
the IKB1 and MICA genes at the centromeric end of the HLA class I
region
JOURNAL Genomics 47 (3), 372-382 (1998)
PUBLISHED 98149985
REFERENCE 2 (bases 1 to 39436)
AUTHORS Shihina,T.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1997) to the DBJ/EMBL/GenBank databases. Takashi
Shihina, Tokai University School of Medicine, Department of
Molecular Life Sciences; Bohseidai, Isehara, Kanagawa 259-11, Japan
(E-mail:tshihina@is.icc.u-tokai.ac.jp, Tel:0463-93-1121,
Fax:0463-94-8884)
FEATURES
    Location/Qualifiers
        1..39436
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /cell_line="BOLETH"
           /cell_type="immunoresponse cell"
           /chromosome="6"
           /clone="745D12"
           /clone_lib="CEPH YAC"
           /dev_stage="adult"
           /haplotype="A2 B62 CW10 DR4"
           /map="6p21.3"
           /sex="male"
           /sub_clone="cosmid clone:TY3A9"
           /tissue_type="blood"
BASE COUNT 10807 a 9318 c 8837 g 10474 t
ORIGIN
Query Match 80.0%; Score 16; DB 22; Length 39436;
Best Local Similarity 94.4%; Pred. No. 1.29e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1376 TCCTTTRAGGGAAGTTC 1393
||||| |||||||
Cp 18 tccttttaggaagtcc 1

RESULT 5
LOCUS AC004184 44706 bp DNA HTG 23-FEB-1998
DEFINITION *** SEQUENCING IN PROGRESS *** Homo sapiens Cosmid R3A; HTGS phase
2, 1 ordered pieces.
ACCESSION AC004184
NID 92905876
KEYWORDS HTG; HTGS_PHASE2.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 44706)
AUTHORS Guillaudeau,T., Janer,M., Wong,G.K.-S., Spies,T. and Geraghty,D.E.
TITLE The complete genomic sequence of 424,015 bp at the centromeric end
of the HLA class I region; gene content and polymorphism
JOURNAL Proc. Natl. Acad. Sci (1998) In press

```

```

REMARK Fred Hutchinson Cancer Research Center
The Clinical Research Division
1100 Fairview Ave. N., P.O. Box 19024
Seattle, WA 98109-1024

REFERENCE 2 (bases 1 to 44706)
AUTHORS Geraghty,D.E. and Olson,M.V.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1998) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
University of Washington Human Genome Center

COMMENT Box 352145 Seattle, WA 98195
Contact: Daniel E. Geraghty (geraghty@fhcrc.org). ***
*** WARNING: Phase 2 High Throughput Genome Sequence ***
*** This sequence is unfinished. It consists of 1 contigs for
* which the order is known. The lengths of the gaps have been
* estimated by the submitter but are not known exactly. When
* sequencing is complete, the sequence data presented in this
* record will be replaced by a single finished sequence
* with the same accession number.
* 1 44706: contig of 44706 bp in length.
FEATURES
    Location/Qualifiers
        1..44706
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /clone="UMGC:R3A"
           /chromosome="6"
           /map="p21"
BASE COUNT 12253 a 9937 c 10116 g 12400 t
ORIGIN
Query Match 80.0%; Score 16; DB 14; Length 44706;
Best Local Similarity 94.4%; Pred. No. 1.29e+01;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 28158 GGAACCTCCCTTAGGGA 28175
||||| |||||||
Cp 1 ggaacttccttaaggga 18

RESULT 6
LOCUS AC004048 102000 bp DNA HTG 28-JAN-1998
DEFINITION *** SEQUENCING IN PROGRESS *** Homo sapiens chromosome 4, BAC clone
B176A5; HTGS phase 1, 6 unordered pieces.
ACCESSION AC004048
NID 92815543
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 102000)
AUTHORS Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 102000)
AUTHORS Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-1998) Department of Genetics, Stanford Human
Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA
*** WARNING: Phase 1 High Throughput Genome Sequence ***
*** This sequence is unfinished. It consists of 6 contigs for
* which the order is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
* as runs of N as a convenience only. When sequencing is complete,
* the sequence data presented in this record will be replaced

```

* 1 2338: contig of 2358 bp in length
* 2359: gap of unknown length
* 2421: contig of 2115 bp in length
* 4536: contig of 2115 bp in length
* 4599: gap of unknown length
* 4600: contig of 2557 bp in length
* 7156: contig of 2557 bp in length
* 7219: gap of unknown length
* 9455: contig of 2236 bp in length
* 9518: gap of unknown length
* 12827: contig of 3308 bp in length
* 12827: contig of 3308 bp in length
* 15116: contig of 2227 bp in length
* 15179: gap of unknown length
* 17217: contig of 2038 bp in length
* 17280: gap of unknown length
* 17281: contig of 2646 bp in length
* 19927: gap of unknown length
* 19989: gap of unknown length
* 23150: contig of 3161 bp in length
* 23213: gap of unknown length
* 26654: contig of 3441 bp in length
* 26717: gap of unknown length
* 31531: contig of 4814 bp in length
* 31594: gap of unknown length
* 31595: contig of 4561 bp in length
* 36218: gap of unknown length
* 39537: contig of 3319 bp in length
* 39600: gap of unknown length
* 47908: contig of 8308 bp in length
* 47971: gap of unknown length
* 51800: contig of 3919 bp in length
* 51953: gap of unknown length
* 60194: contig of 8241 bp in length
* 60257: gap of unknown length
* 72656: contig of 12399 bp in length
* 72719: gap of unknown length
* 81871: contig of 9152 bp in length
* 81934: gap of unknown length
* 96177: contig of 14243 bp in length
* 96178: gap of unknown length
* 96240: contig of 13606 bp in length
* 109847: gap of unknown length
* 109910: contig of 11902 bp in length.

FEATURES
source
1. .121811
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="bd3-6"
/map="16"

BASE COUNT 30294 a 30125 c 29106 g 30970 t 1316 others

Query Match 85.0%; Score 17; DB 14; Length 121811;
Best Local Similarity 94.7%; Pred. No. 2.89e+00;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 117521 GAACCTCCCTAAAGGAGG 117539
||||| |||||||||
Qy 2 gaacttcctaaaggagg 20

RESULT 2 128278 215 bp DNA PAT 30-OCT-1996
LOCUS

DEFINITION Sequence 5 from patent US 5569830.
ACCESSION 128278
NID g1819054
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 215)
Bennett,A., Labavitch,J.M., Powell,A. and Stotz,H.
Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;
FEATURES Location/Qualifiers
source 1. .215
/organism="unknown"
BASE COUNT 15 a 8 c 25 g 26 t 141 others
ORIGIN

Query Match 80.0%; Score 16; DB 18; Length 215;
Best Local Similarity 35.3%; Pred. No. 1.29e+01;
Matches 6; Conservative 10; Mismatches 1; Indels 0; Gaps 0;

Db 78 THYTTNNSGADSKTV 94
|:::|:::|:::|
Cp 18 tccttttaggaagttc 2

RESULT 3 XLXWNT5C 1188 bp RNA VRT 05-JUL-1993
LOCUS

DEFINITION X.laevis Xwnt-5C mRNA.
ACCESSION X73510
NID g313267
KEYWORDS Wnt/wingless gene; Xwnt-5C gene.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae;
Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1188)
Koster,J.G., Kuiken,G.A., Stegeman,B., Peterson,J., Elizema,K.,
Stabel,L., Dekker,E.J. and Destre,O.H.J.
Differential Xwnt-5C expression during early development of Xenopus
laevis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1188)
Koster,J.G.
Direct Submission
TITLE Submitted (22-JUN-1993) J.G. Koster, Hubrecht Laboratory,
Uppsalaalaan 8, NL 3584 CT Utrecht, NETHERLANDS

FEATURES Location/Qualifiers
source 1. .1188
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/dev_stage="stage 17 embryo"
40. .1122
/gene="Xwnt-5c"
40. .1122
/gene="Xwnt-5c"
/codon_start=1
/db_xref="PID:g313268"
/db_xref="SWISS-PROT:P33945"
/translation="WTPILRLILLSSLLSCWKQSVGVGANSWMSALNPVORPEMETIG
APLCSQTLGSLPGKQKCOLQYDMVHIGEGAKTGKCCQHFHRKNCSTVDNNS
VFGVMOIGREASFTYAISSAGVYNAISRACREGELSTCGCSRTPRPKDLPLDWLWG
GCGDNVEYGFKEFVDARKEFKPKGSEQLNMLQNNRAGRAVYKLDVAC
KCHGVSGSLKTLQLADFRKVGVEYIKEDYSAASMLNKRNLKLEQVQNFNPPTG
EDLVLDPSDPYCLYNETGSLGTHGRQCNKTSEGMDCGELCCCGRGYDQFTVQVER
CHCKEQWCCFVCKECTEIVDQVCK"

BASE COUNT 285 a 287 c 359 g 257 t
ORIGIN

Query Match 80.0%; Score 16; DB 17; Length 1188;
Best Local Similarity 100.0%; Pred. No. 1.29e+01;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 591 GAACCTCCCTAAAGG 606
||||| |||||||||
Qy 2 gaacttcctaaagg 17

RESULT 4 AB000879 39436 bp DNA PRI 04-APR-1998
LOCUS

MORF (TM)

Release 3.0.5AA John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1997 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
on: Wed May 27 02:58:29 1998; Maspar time 53.56 Seconds
623.063 Million cell updates/sec
Tubular output not generated.

Title: >SEQ2
Description: (1-20) from new.seq
Perfect Score: 20
N.A. Sequence: 1 ggaacttcctaaaggagg 20
Comp: ccttgaagggtttccctcc

Scoring table: TABLE default
Gap 10

Nmatch STD : Dbase 0; Query 0
Searched: 457423 seqs, 834342348 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb154
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro 12:em_v1
Database: genbank106
13:gb_ba 14:gb_htg 15:gb_in 16:gb_om 17:gb_ov 18:gb_pat
19:gb_ph 20:gb_pl 21:gb_pr1 22:gb_pr2 23:gb_ro 24:gb_st
25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 6.503; Variance 2.805; scale 2.318
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	17	85.0	121811 14 AC000096 *** SEQUENCING IN PROG 2.89e+00
2	16	80.0	215 18 128278 Sequence 5 from patent 1.29e+01
3	16	80.0	1198 17 XLXNTSC X.laavis Xwnt-5C mRNA. 1.29e+01
4	16	80.0	39436 22 AB000879 Homo sapiens genomic D 1.29e+01
5	16	80.0	44706 14 AC004184 *** SEQUENCING IN PROG 1.29e+01
6	16	80.0	102000 14 AC004048 *** SEQUENCING IN PROG 1.29e+01
7	16	80.0	148960 22 AB000882 Homo sapiens genomic D 1.29e+01
8	16	80.0	267156 22 U66059 Human germline T-cell 1.29e+01
9	15	75.0	533 20 BB58SRR B.bronchiariti 5.8S rR 5.48e+01
10	15	75.0	537 20 BT58SRR B.tenella 5.8S rRNA ge 5.48e+01
11	15	75.0	557 27 SVU27968 Stealth virus clone C1 5.48e+01
12	15	75.0	1398 17 GAX1EX1 G.gallus axonin-1 gene 5.48e+01
13	15	75.0	1416 13 ASU78028 Arthrobacter sp. beta- 5.48e+01
14	15	75.0	3304 22 HSL81905 Homo sapiens (subclone 5.48e+01
15	15	75.0	3560 13 HECURES Helicobacter pylori ur 5.48e+01

C	16	15	75.0	3616 18 A24198 H.pylori ureE, uref, u 5.48e+01
	17	15	75.0	14030 13 HPAA000528 Helicobacter pylori se 5.48e+01
	18	15	75.0	38017 20 AB004534 Schizosaccharomyces po 5.48e+01
C	19	15	75.0	41623 22 HSN69P4 Human DNA sequence fro 5.48e+01
	20	15	75.0	42759 20 AB004535 Schizosaccharomyces po 5.48e+01
C	21	15	75.0	79611 22 AC002486 Human BAC clone RG3670 5.48e+01
C	22	15	75.0	81674 22 HS722E9 Homo sapiens DNA sequ 5.48e+01
C	23	15	75.0	81826 22 HSU78045 Human collagenase and 5.48e+01
C	24	15	75.0	86156 22 HUAC002550 Human Chromosome 16 BA 5.48e+01
C	25	15	75.0	128669 14 HS232624 Human DNA sequence *** 5.48e+01
	26	15	75.0	135038 22 HUM1WD703 Homo sapiens ADP/ATP c 5.48e+01
C	27	15	75.0	137527 14 HS25J6 Human DNA sequence *** 5.48e+01
C	28	15	75.0	266485 14 HS212A2 S.lividus gene for l6S 2.20e+02
	29	14	70.0	267 13 SL16SRRN1 Shewanella alga DNA gy 2.20e+02
C	30	14	70.0	1257 13 AF005687 Sequence 1 from patent 2.20e+02
	31	14	70.0	1404 18 I62881 M.sativa MMK4 mRNA for 2.20e+02
C	32	14	70.0	1695 20 MSMMK4 Hendra virus structure 2.20e+02
C	33	14	70.0	2698 27 AF010304 H.sapiens germline gen 2.20e+02
C	34	14	70.0	2908 21 HSIGK15 C.fusiformis plasmid p 2.20e+02
C	35	14	70.0	4079 20 CFCF2 Helicobacter pylori se 2.20e+02
C	36	14	70.0	10433 13 HPAA000618 Methanobacterium therm 2.20e+02
	37	14	70.0	10818 13 AE000829 Border disease virus s 2.20e+02
	38	14	70.0	12255 27 AF002227 Methanococcus jannasch 2.20e+02
	39	14	70.0	12284 13 U67467 Rattus norvegicus neur 2.20e+02
	40	14	70.0	12728 23 RNU22062 Aquifex aeolicus secti 2.20e+02
C	41	14	70.0	15015 13 AE000774 Caenorhabditis elegans 2.20e+02
	42	14	70.0	43245 15 CEM163 Human BAC clone RG05F 2.20e+02
C	43	14	70.0	78809 22 HSAC000056 Human DNA sequence *** 2.20e+02
	44	14	70.0	141535 14 HS327J16 Arabidopsis thaliana D 2.20e+02
	45	14	70.0	200611 20 ATFC40 Arabidopsis thaliana D 2.20e+02

ALIGNMENTS

RESULT LOCUS	1	AC000096	121811 bp	DNA	HTG	30-JAN-1998
DEFINITION	*** SEQUENCING IN PROGRESS *** Mus musculus Chromosome 16 BAC Clone bd3-6 Syntenic to Homo sapiens 22q11.2 DGR Region; HTGS phase 1, 21 unordered pieces.					
ACCESSION	AC000096					
NID	g2826455					
KEYWORDS	HTG: HTGS, PHASE1.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
AUTHORS	1 (bases 1 to 121811) Galili,N., Baldwin,S., Lund,J., Reeves,R., Gong,W., Wang,Z., Roe,B.A., Emanuel,B.S., Nayak,S., Mickanin,C., Budarf,M.L. and Buck,C.A.					
TITLE	A Region of Mouse Chromosome 16 is Syntenic to the Digeorge, Velo-Cardio-Facial Syndrome Minimal Critical Region					
JOURNAL	Unpublished (1996)					
REMARK	The genes were identified by comparing with human genomic and cDNA sequences and RT-PCR of 12 day post conception mouse embryos total RNA					
REFERENCE	2 (bases 1 to 121811)					
AUTHORS	Roe,B.A.					
TITLE	Direct Submission					
JOURNAL	Submitted (16-NOV-1996) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA					
COMMENT	*** WARNING: Phase 1 High Throughput Genome Sequence *** *** This sequence is unfinished. It consists of 21 contigs for which the order is not known; their order in this record is arbitrary. In some cases, the exact lengths of the gaps between the contigs are also unknown; these gaps are presented as runs of N as a convenience only. When sequencing is complete, the sequence data presented in this record will be replaced by a single finished sequence with the same accession number.					

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☐ BLACK BORDERS

☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES

☐ FADED TEXT OR DRAWING

☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING

☐ SKEWED/SLANTED IMAGES

☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS

☐ GRAY SCALE DOCUMENTS

☐ LINES OR MARKS ON ORIGINAL DOCUMENT

☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY

☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)